مبادئ في الإحصاء والاحتمالات

د. حميد عويِّد مشرف العكله

أستاذ في قسم العلوم الأساسية- عمادة السنة الأولى المشتركة

د. إبراهيم عبد العزيز إبراهيم الواصل

أستاذ في قسم الإحصاء وبحوث العمليات- كلية العلوم

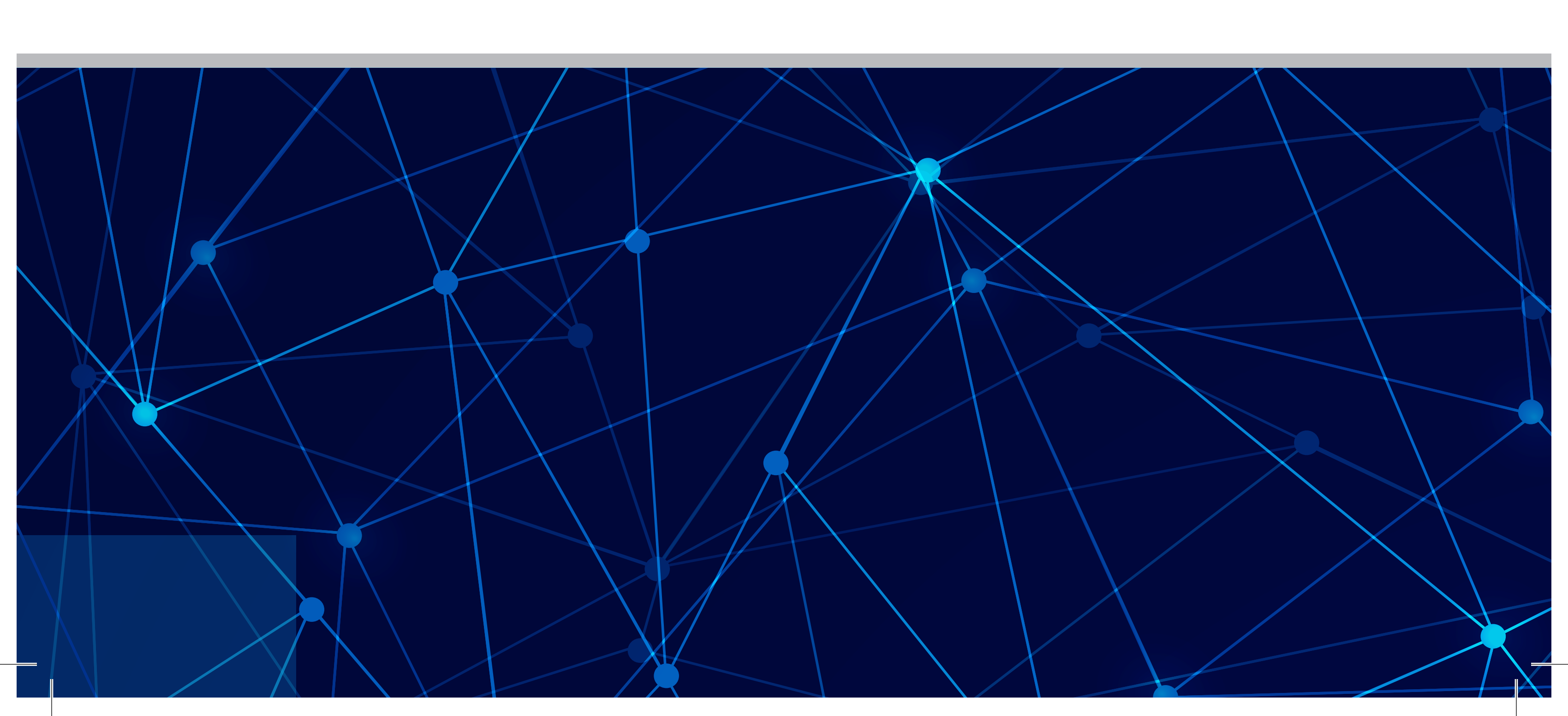
د. منصور محمد علي شراحيلي

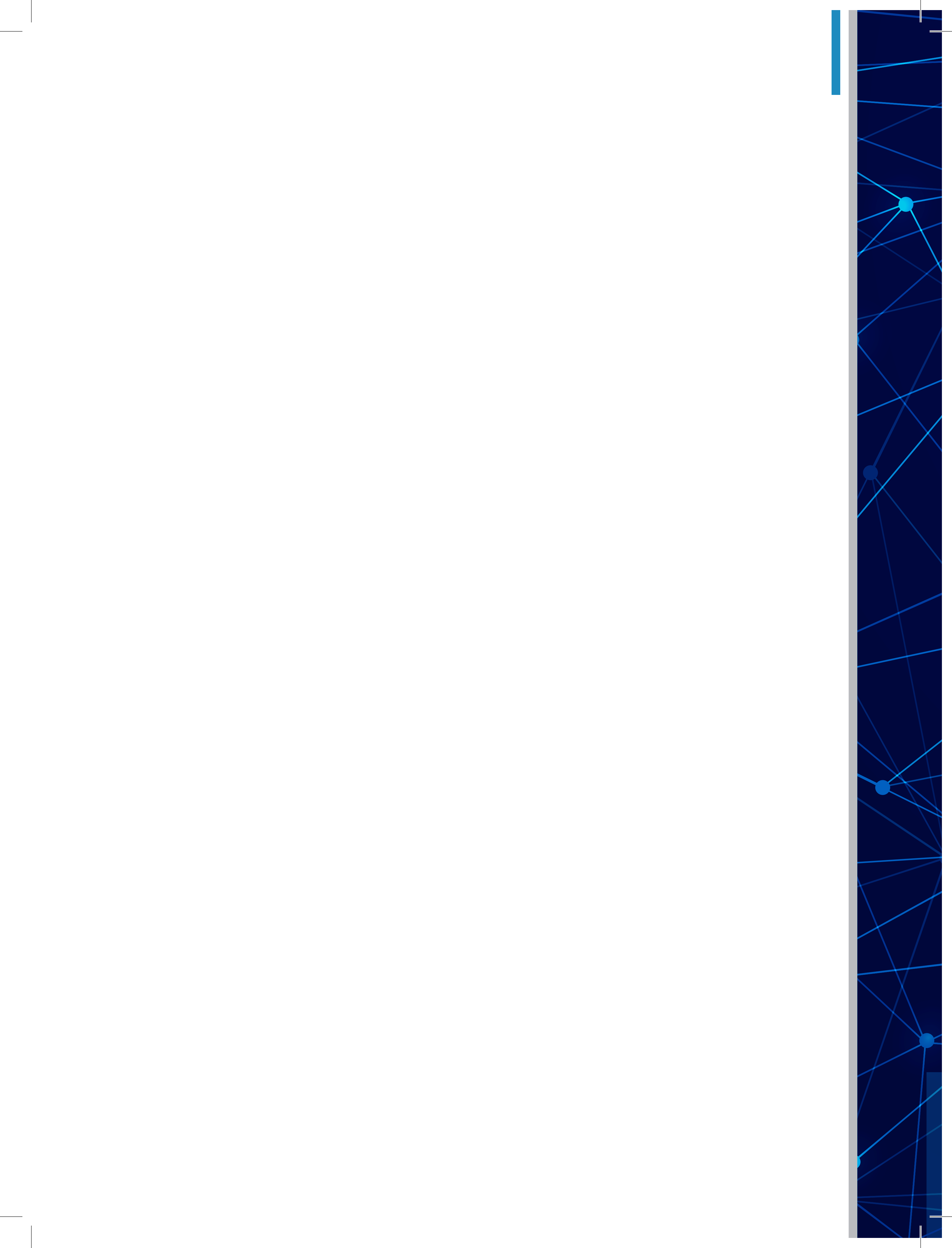
أستاذ مساعد في قسم الإحصاء وبحوث العمليات- كلية العلوم

د. إبراهيم علي حسن النفيسه

أستاذ مساعد في قسم الإحصاء وبحوث العمليات- كلية العلوم

KING SAUD UNIVERSITY





محتوى الكتاب



مقدمة الكتاب

الحمد للَّه ربِّ العالمين، والصلاة والسَّلام على نبينا محمّد وعلى آله وأصحابه ومن اتَّبعهم بإحسان إلى يوم الدّين. أمّا بعد، فقد أوصانا اللَّه تعالى بطلب العلم وحثّنا على اكتشاف أسرار مخلوقاته وسبر مكنوناتها، ولكن بالعلم والمعرفة وتقديم الدّليل والحجّة الدَّاحضة، حيث يقول ربّنا عزَّ وجل في سورة الرحمن (يَا مَعْشَرَ الْجِنِّ وَالْإِنْسِ إِنِ اسْتَطَعْتُمْ أَنْ تَنْفُذُوا مِنْ أَقْطَارِ السَّمَاوَاتِ وَالْأَرْضِ فَانْفُذُوا لَا تَنْفُذُونَ إِلَّا بِسُلْطَانٍ) (الآية 33)، وذكرنا ربّنا تعالى أنَّ ما لدينا من علوم ما هو إلاَّ بالشيء القليل حيث يقول سبحانه وتعالى في الآية 85 من سورة الإسراء (وَمَا أُوتِيتُمْ مِنَ الْعِلْمِ إِلَّا قَلِيلا).

إنَّ علم العشوائيّات (الإحصاء الرياضي ونظرية الاحتمالات) يُعدّ من العلوم التّطبيقيّة الأكثر انتشاراً في الوقت المُعاصر، وذلك لأنَّ هذا العلم قد أصبح على علاقةٍ وثيقة بمعظّم فروع العلوم العلميّة والانسانيّة، حيث يُلاحظ أنَّ الكثير من الاختصاصات تحتاج إلى تحاليل وعروض واختبارات إحصائيّة من أجل تثبيت صحة نتائج دراساتها. أمّا عن تاريخ علم الاحصاء فليس هناك ما يشير إلى زمن ولادته بشكلٍ دقيق، فالبعض يَذكر أنَّه كان حوالي عام 1662، والبعض الآخر يذكر أنَّه يعود إلى حوالي عام 1749. لكن في الواقع يُعتقد أنَّ علم الإحصاء قديم قِدَم تاريخ الحساب عندما بدأت المراحل الأولى للتّجمعات البشريّة حيث سادّت فيها مظاهر السلطة وحبّ التملُّك، وظهور العدِّ والتصنيف لحصر الممتلكات من الأنعام والجُند والجِنَان، وذلك لأنَّ عمليّة العدّ مع التصنيف ما هي إلاَّ عمل إحصائيّ ابتدائيّ. لذلك يُنظر إلى الإحصاء على أنَّه توأم الحساب أو ربما كان أحدهما هو الشقيق الأكبر للآخر. أمّا استخدامه كمفردةٍ لغويّة، فإنَّه كان متداولاً بين النَّاس منذ زمنٍ بعيدٍ جداً والدليل على ذلك ورود كلمة «إحصاء» في القرآن الكريم مرّاتٍ عديدة، وهذا يعني أنَّ كلمة «إحصاء» كانت ذاتَ دلالةٍ واضحةٍ منذ أمدٍ بعيد. إنَّ أول كتاب نُشر في الإحصاء كان عام 1845 من قبل الرّياضيّاتي (الاكتواريّ) الإنكليزي نيسون Francis Gustavus Paulus Neison وحمل عنوان «مساهمات في الإحصاء الحيويّ Contributions to Vital Statistics». لقد نَشِطَ في هذا المجال عددٌ كبيرٌ جداً من المهتّمين بهذا العلم وأكثرهم من علماء العلوم الطّبيعيّة (الرّياضيات، الإحصاء، الفيزياء وعلم الحياة)، وكان من أبـرّز هؤلاء: بـييز (1702-1761) Thomas Bayes، بواسـون (1781-1840) Siméon Denis Poisson، بيرسـون (1857-1936) Karl Pearson وفيشر (1890-1962) Ronald Fisher.

أمـَّا علم الاحتمالات فهو حديث العهد (نسبياً) إذا ما قورن بعلم الإحصاء، فقد نشأ الحساب الاحتمالي من خلال محاولة إيجاد الحلول لبعض ألعاب الحظ والمراهنات. هذا وقد حظيت الاحتمالات باهتمامٍ كبيرٍ من قِبَل علماء العلوم الطبيعية عامّةً وعلماء الرّياضيّات خاصّةً، وذلك لما لهذا العلم من تطبيقاتٍ مهمّة في معظم مجالات العلوم التّطبيقيّة والنّظريّة على حدٍ سواء، فعلى سبيل المثال لا الحصر يمكن القول إنَّ نّظرية الاحتمالات أصبحت الحجرِ الأساس لمعظم الدّراسات المهتمّة بالتّجارب والظّواهر العشوائيّة ونمذجتها. إنَّ أول كتاب نُشر في الحساب الاحتمالي كان عام 1657 من قبل هويغنز (1629-1695) Christiaan Huygens وحمل عنوان «حسابات لألعاب الحظ». إنَّ عام 1713 يعدُّ التّاريخ الحقيقي لولادة علم الاحتمالات بعد أن تمَّ نشر أول مَقالةٍ علميَّةٍ متقّنة للرّياضيّاتي السويسري يعقوب برنولي (1655-1705) Jakob Bernoulli (قدِّمت من قبل أخيه بعد ثمان سنواتٍ من وفاته). أمـَّا النّاشطون في هذا المجال فهم كثـيرون جداً أيضاً، ومن أبـرزهم: هويغنز، يعقوب برنولي والرّياضيّاتي الروسيّ كلموغوراف (1903-1987) Andrey Kolmogorov.

في إطار تطوير العمل الأكاديمي من أجل الوصول إلى مستوىً علميٍّ مرموق ينهل منه الطّلبة في الجامعات فقد تمّ بفضلٍ من الله تعالى وعونه إعداد هذا الكتاب ليكون لبِنَةً من لبِنَات الصّرح العلمي للجامعات والمكتبات العربيّة. لقد تمَّ إعداد هذا الكتاب بحيث يكون متناسباً مع السويّةِ العلميّةِ للطّلبةِ الجامعيين في العلوم الإنسانيّة ومنسجماً مع خطة تدريسيّة لفصلٍ دراسيٍّ واحد بمعدّل ثلاث ساعاتٍ معتمدة.

في الواقع لم يكن إنجاز هذا الكتاب سهلاً لأنَّه يتعلّق بتقديم محتوىً علميٍّ لطّالبٍ في مسار علوم إنسانيّة، فقد حاولنا جاهدين إغناء موضوعات هذا الكتاب بقدرٍ جيّدٍ من المعارف العامّة في الإحصاء والاحتمالات التي تناسب طلبة المسار الإنسانيّ في أيّة جامعة وبحيث تلبّي الكثير من احتياجاته الإحصائيّة خلال مسيرته التعليميّة الجامعيّة. كما حاولنا عرض النّصوص والصيغ بشكل يمكن للقارئ تفهمها دون جهدٍ أو عناءٍ وموضّحين ذلك بالجداول والرّسوم والعروض البيانيّة عند الضرورة.

بالطّبع لدى تقديم المعلومات في هذا الكتاب تمّ الأخذ بالحسبان المستوى العلمي والمعرفي للطّالب الجامعيّ الدّارس في مسار علم إنسانيّ، ولذلك بدأنا بمعلوماتٍ بسيطةٍ متوفرٌ بعضها لدى الطّالب من مرحلة التعليم قبل الجامعيّ، ومن ثمّ التدرّج في رفع مستوى المعلومة إلى أن بلغت ذروتها في دراسة الاحتمالات. كذلك حاولنا قدر المستطاع إظهار التّرابط العلميّ والموضوعيّ للفقرات المقدَّمة بحيث تبدو الفقرات للطالب متسلسلة بشكلٍ منطقيّ، بمعنى أنَّ المعلومات المتراكبة بُنيت (في معظم الحالات) على نحوٍ متصاعد.

لقد قدَّمنا في هذا الكتاب سبعة فصولٍ تناولنا في طيّاتها موضوعات أساسية في الإحصاء والاحتمالات، وكانت محتوياتها على النحو الآتي:

- **الفصل التّمهيدي**، وقد قدِّم فيه بعض المفاهيم الريّاضياتيّة اللّازمة لدراسةِ الفصول اللاحقة.
- **الفصل الأول**، وقد خُصِّصَ لتنظيم البيانات الإحصائيّة وعرضها بيانياً.
- **الفصل الثّاني**، وقد خُصِّصَ لتقديم بعض مقاييس الموضع للبيانات الإحصائيّة.
- **الفصل الثّالث**، وقد خُصِّصَ لتقديم بعض مقاييس الاختلاف للبيانات الإحصائيّة.
- **الفصل الرّابع**، وقد خُصِّصَ لتقديم مبادئ أوليّة في الارتباط والانحدار الخطّي.
- **الفصل الخامس**، وقد خُصِّصَ لشرح مفهوم التّجارب العشوائيّة وحساب احتمالات حوادث متعلّقةٍ بها.
- **الفصل السّادس**، وقد خُصِّصَ لتقديم مفهوم المتغيِّر العشوائيّ وعرض بعض نماذجها البسيطة والمهمّة.

لقد قمنا بإغناء موضوعات وفقرات هذا الكتاب بعددٍ لا بأس به من الأمثلة المحلولة، ومن ثمَّ إدراج عددٍ آخر في نهاية كل فصل من التّمارين غير المحلولة المناسبة للفِكَر المقدَّمة في الفقرات النّظريّة. كما قمنا بترقيم الفقرات المهمّة والعلاقات والجداول والأشكال بطريقة تُسهِّل على الطالب التّنقّل فيما بينها، وأخيراً ننوِّه إلى أنَّنا قمنا بتشكيل الكثير من الكلمات وتجاوزنا عن البعض الآخر بشّكل متعمَّد من أجل أن يتعوّد الطّالب على القراءة الصحيحة للكلمة وإذا ما شُكِّلت الكلمة لاحقاً فهو من باب التذكير فقط.

في الحقيقة إنَّ ما قُدِّمَ في طيّات هذا الكتاب ليس إلاَّ غيض من فيض، ولذلك سيجد القارئ الكثير من الاقتراحات على إضافة فقرةٍ هنا وحذف أخرى هناك وتعديل على فقرةٍ في موضعٍ آخر، وهذا أمر قلّما يَسلم منه مُؤلَّف مهما بُذِل من جهد لأجله. لذلك نودّ أن نعرب عن شكرنا وتقديرنا العميقين لكل من يُبدي لنا ملاحظةً مفيدةً أو نقداً بنّاءً حول هذا الكتاب آملين من اللّه تعالى أن نكون قد وفّقنا في تقديمه بالشّكل المناسب والمفيد.

قبل أنّ نختم مقدِّمتنا هذه ننوّه إلى أنّنا اعتمدنا في التّعريب على بعض المعاجم العربيّة ومنها معجم الرّياضيات الصّادر عن مؤسسة الكويت للتّقدُّم العلميّ، وكذلك معجم الرّياضيات من منشورات دار النّشر أكاديميا في بيروت.

اللّهم انفعنا بما علمتنا وعلِّمنا ما ينفعنا وزدنا علماً وعملاً متقبلاً إنَّك أنت السميع العليم

الرياض: يوم الخميس في 1438/08/08 الموافق لـــ 04/ 05/ 2017

المؤلِّفون

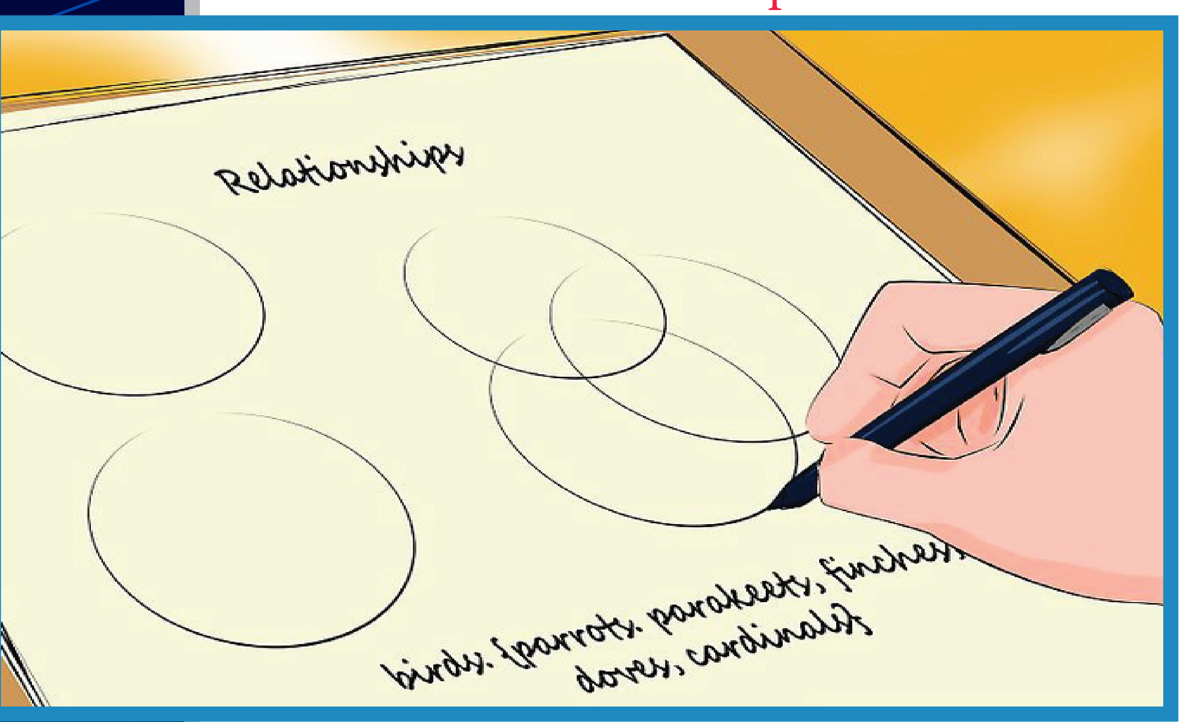
شكر وتقدير

يودّ المؤلِّفون تقديم الشّكر الجزيل إلى عميد السّنة الأولى المشتركة الدكتور نامي الجهني ووكيله للشؤون الأكاديمية الدكتور عبدالمجيد الجريوي، وإلى الدكتور ناصر التركي رئيس قسم العلوم الأساسيّة في عمادة السّنة الأولى المشتركة في جامعة الملك سعود على رعايتهم دعم تأليف هذا الكتاب، كما يشكرون جامعة الملك سعود التي أسهمت في طباعة هذا الكتاب ونشره، وشكرهم موصول إلى منسِّق ومصمِّم ومنتج رسومات هذا الكتاب الأستاذ فادي حسن.

الفصل التمهيدي

مفاهيم أساسية من الرياضيات

Basic Concepts of Mathematics



المقدمة:

نقدِّم في هذا الفصل بعض المفاهيم الأساسيّة في الرّياضيات علماً أنَّ الطالب الجامعي قد تعرَّف على الكثير منها خلال مراحل تعليمه السّابقة، ولكن لابدَّ من هذا التقديم لتذكير القارئ ببعض المفاهيم الرّياضياتيّة الأساسيّة التي سيحتاجها هذا الكتاب المخصَّص لتدريس مقرَّر مبادىء في الإحصاء والاحتمالات من مقرَّرات السّنة الأولى المشتركة. إنَّ البراهين والإثباتات الخاصة بفقرات هذا الفصل سيتمّ تجاوز معظمها حيث يمكن لمن يودّ الاطلاع عليها الرجوع إلى المراجع ذات الصلة والتي ذكر بعضها في نهاية هذا الكتاب.

- ت - ١ - المجموعات وخصائصها
- ت - ٢ - الدوال الحقيقية

ت ١
المجموعات وخصائصها

إنّ مفهوم المجموعة يُعدُّ من الأهمية بمكان في مجال الرّياضيات عامة ذلك أنّه يأتي في المرتبة الثّانيّة من حيث الأهمية بعد مفهوم الأعداد، وقد أسِّس على يد الرّياضياتي الألماني **كانتور** Georg Cantor (1845-1918)، ولكنَّنا سنقدِّم قسماً يسيراً من نظرية المجموعات مع التركيز على المواضيع التي تهمّنا في إطار هذا الكتاب، وسنستخدم ما اتفق عليه علمياً فقط.

في الحقيقة إنّ نظرية المجموعات هي أحد فروع علم المنطق الرّياضياتي الذي يهتمّ بدراسة المجموعات والعمليات الجبرية عليها، فبالرغم من أنّ أي تجمّع لأشياء ذات طبيعة ما يمكن تمثيلها في مجموعة، إلاّ أنّنا سنقوم بتطبيق نظرية المجموعات في معظم الأحيان على الأشياء التي لها صلة بعلم الرّياضيّات (**وكجزء منها الاحتمالات والإحصاء**)، وذلك أنّه يمكن للغة نظرية المجموعات أن تستخدم في صياغة التعاريف لمفاهيم رياضياتيّة. إن مفهوم المجموعة الذي سيقدِّمه التّعريف الآتي يعود إلى كانتور.

> **ت-١-١- تعريف (المجموعة) (Set)**
>
> إن المجموعة هي أي تجمّع لأشياء (**أو كائنات**) محدَّدة ومتمايزة جيداً (**أو مختلفة بعضها عن البعض الآخر**) بحسب حدسنا أو تفكيرنا، وصُبَّت في قالب موحّد (**أو في شكل كلي واحد**).

بمعنى آخر. المجموعة هي **تَجمُّع** لأشياء (**أو كائنات**) محدَّدة ومتمايزة بشكل جيد وليست مرتَّبة بالضرورة، ولكنَّها تشترك فيما بينها بصفة واحدة على الأقل، ويُعبَّر عنها رياضياتياً بقوسين كبيرين متقابلين يحتويان بينهما الرموز الممثِّلة لتلك الأشياء (**أو الكائنات**) وذلك على الشكل $\{\bullet,\bullet,\ldots,\bullet\}$، فعلى سبيل المثال تُمثَّل مجموعة الأعداد في النظام العشري كما يلي:

$$\left\{0,1,2,3,4,5,6,7,8,9\right\}$$

ويُدعى هذا العرض بـ **التمثيل السَّردي** Enumerative Representation للمجموعة، ولكن توجد طرائق أخرى لعرض المجموعات، ومنها **طريقة القاعدة** Bar Notation، وفي هذه الطريقة يُرمز للصفة المميِّزة لعناصر المجموعة بحرفٍ مع التنويه إلى دلالة هذا الحرف ومن ثمَّ يحاط الجميع بقوسين كبيرين، فعلى سبيل المثال لو أخذنا طلاب السّنة الأولى المشتركة (**في عام معيَّن**) فإنَّهم يكوِّنون مجموعة (**ولتكن E مثلاً**) حيث يوجد من الجنسين الذكور والإناث، ولكنَّهم جميعاً تجمع بينهم صفة الدّراسة في السّنة الأولى المشتركة، وحينئذٍ يمكننا أن نكتب:

$\{x \mid x$ **طالب في السّنة الأولى المشتركة**$\}$

وهنا نشير إلى أنَّ الرّمز (|) داخل قوسي المجموعة يقصد به "علماً أنَّ" أو "حيث أنَّ".

ت-١-٢- ملاحظات

١- لقد درجت العادة على أن يُرمز للمجموعة بأحرف لاتينيّة كبيرة من قبيل A و B و C أو....
أو بأحرف إغريقيّة كبيرة من قبيل Ω و Ξ و Θ و $\cdots$.

٢- إنَّ الأشياء (أو الكائنات) المكوّنة للمجموعة تُدعى **عناصر** Elements ويُرمز لها عادةً بأحرف لاتينيّة صغيرة من قبيل a و b و c و... ، أو بأحرف إغريقية صغيرة من قبيل α و β و γ و $\cdots$ أو برموزٍ أخرى حسب طبيعة العناصر المكوِّنة للمجموعة. من جهة أخرى، فإذا كان a عنصراً من مجموعة A فعندئذٍ يُقال إنَّ العنصر a ينتمي إلى A ويُرمز لذلك بالشكل $a \in A$، وأمّا إن كان a ليس عنصراً من المجموعة A فحينئذٍ يُرمز لذلك بالشكل $a \notin A$.

٣- يُقال عن المجموعة التي لا تملك أي عنصر إنَّها مجموعة خالية، ويرمز لها بـ $\varnothing$، أي أنَّ:

$$\varnothing = \{\ \}$$

٤- يُقال عن مجموعتين A و B غير خاليتين إنَّهما **متساويتان** إذا كانتا تملكان العناصر نفسها، ويكتب حينئذٍ $A = B$، فعلى سبيل المثال:

أ- نجد أنَّ المجموعتين $A = \{7, 2, 5, 1, 15, 3\}$ و $B = \{1, 2, 3, 5, 7, 15\}$ متساويتان لأنَّهما تملكان العناصر نفسها.

ب- إنَّ المجموعة $A = \{a, b, c, \ldots, x, y, z\}$ تساوي المجموعة الآتيّة لأنَّ لهما العناصر نفسها.

$$B = \{x \mid x \text{ حرف أبجدي إنكليزي}\}$$

٥- يُقال عن مجموعة A إنَّها **مجموعة جزئيّة** من مجموعة B إذا كانت جميع عناصر المجموعة A تنتمي إلى المجموعة B ويُكتب حينئذٍ $A \subseteq B$، وفي حال وجود عنصر واحد على الأقل في B لا ينتمي إلى A فحينئذٍ يُقال إنَّ المجموعة A **محتواة تماماً** في المجموعة B ويُكتب $A \subset B$.

٦- إذا كنّا نتعامل مع مجموعات جزئيّة من مجموعة ما غير خالية Ω، فعندئذٍ يُقال عن Ω إنّها **مجموعة شاملة**.

٧- إنَّ كلّ مجموعة محتواه في نفسها. أي أنَّه من أجل مجموعة A يمكننا أن نكتب $A \subseteq A$.

٨- إذا كانت A وB مجموعتين من مجموعة Ω، وكان $A \subseteq B$ و $B \subseteq A$، فعندئذٍ تكون المجموعتان متساويتين، ويُكتب $A = B$.

٩- نشير هنا إلى أنَّه من أجل مجموعة غير خالية Ω تكون المجموعة الخالية $\varnothing$ محتواه فيها وفي أية مجموعة جزئيّة منها، بمعنى أنَّه من أجل $\varnothing \neq A \subseteq \Omega$ يكون لدينا $\varnothing \subset A$ أيضاً.

ت-١-٣- أمثلة

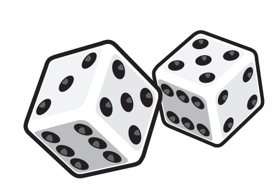
١- لنقم برمي حجر نرد (سداسي الوجوه) لمرَّةٍ واحدةٍ فقط، فعندئذٍ سنحصل على أحد الأعداد الطبيعيّة 1 أو 2 أو... أو 6، علماً أنَّ مجموعة الأعداد الطبيعيّة هي 1 و 2 و 3 و...، ويُرمز لها بـ $\mathbb{N}$.



من هذه العملية يكون لمجموعة القيم (ولتكن A مثلاً) العرض الآتي:

$$A = \{1,2,3,4,5,6\}$$

٢- لنقم برمي حجر نرد لمرَّتين متتاليتين، ولنعيِّن المجموعة الناتجة عن فرق القيمة التي ستظهر في المرّة الثّانيّة من القيمة التي ظهرت في المرّة الأولى.



من هذه العملية نحصل على أحد الأعداد الصحيحة 5- أو 4- أو... أو 1- أو 0 أو 1 أو 2 أو... أو 5، علماً أنَّ مجموعة الأعداد الصحيحة هي: ... و 3- و 2- و 1- و 0 و 1 و 2 و 3 و...، ويُرمز لها بـ $\mathbb{Z}$.

ومن ثمَّ يكون لمجموعة القيم (ولتكن B مثلاً) العرض الآتي:

$$B = \{-5,-4,-3,-2,-1,0,1,2,3,4,5\}$$

٣- لنقم برمي حجر نرد لمرَّتين متتاليتين، ولنعيِّن المجموعة الناتجة عن قسمة القيمة التي ظهرت في المرّة الأولى على القيمة التي ستظهر في المرّة الثّانيّة.

من هذه العملية سنحصل على مجموعة (ولتكن C مثلاً) تحوي الأعداد النسبيّة (أو الأعداد العاديّة أو الأعداد الكسريّة) الآتيّة:

$$\frac{1}{6},\frac{1}{5},\frac{1}{4},\frac{1}{3},\frac{1}{2},1,\frac{2}{5},\frac{2}{3},2,\frac{3}{5},\frac{3}{4},\frac{3}{2},3,\frac{4}{5},\frac{4}{3},4,\frac{5}{6},\frac{5}{4},\frac{5}{3},\frac{5}{2},5,\frac{6}{5},6$$

ومن ثمَّ سيكون للمجموعة C العرض الآتي:

$$C = \left\{\begin{array}{l}\frac{1}{6},\frac{1}{5},\frac{1}{4},\frac{1}{3},\frac{1}{2},1,\frac{2}{5},\frac{2}{3},2,\frac{3}{5},\frac{3}{4},\frac{3}{2},3,\frac{4}{5},\frac{4}{3},4,\\ \frac{5}{6},\frac{5}{4},\frac{5}{3},\frac{5}{2},5,\frac{6}{5},6\end{array}\right\}$$

علماً أنَّ مجموعة الأعداد النّسبيّة هي مجموعة كلّ الأعداد التي تكتب على شكل نسبة لعددين صحيحين أوليين فيما بينهما (مع الأخذ بالحسبان أنَّ المقام لا يساوي الصفر)، ويُرمز لهذه المجموعة بـ $\mathbb{Q}$، أي أنَّ:

$$\mathbb{Q} := \left\{ x \mid x = \frac{a}{b} \;\; ; a,b \in \mathbb{Z},\, b \neq 0 \text{ and } a,\, b \text{ are coprime} \right\}$$

٤- إنَّ مجموعة الأعداد غير النسبيّة (الأعداد التي لا يمكن كتابتها على شكل نسبة لعددين صحيحين - مع الأخذ بالحسبان أنَّ المقام لا يساوي الصفر- ويُرمز لهذه المجموعة بـ $\mathbb{I}$) يمكن عرضها على النّحو الآتي:

$\mathbb{I} = \{x \mid x$ عدد غير نسبي $\}$

٥- لنقم بقذف قطعة نقود معدنية لمرَّةٍ واحدةٍ فقط، فعندئذٍ يكون لمجموعة الرموز (ولتكن D مثلاً) التي تنتج عن هذه العملية هي:

$D = \{H, T\}$



الصورة *Head* الشعار *Tail*

علماً أنَّ الحرفين H وT يشيران إلى كلمة Head وTail على الترتيب. إنَّ سبب استخدام Head تعود إلى القطع النقدية المعدنية التي استُخْدِمت في دراسة هذا النوع من التجارب قديماً حيث كان ينقش على أحد وجهيها صورة رأس أو جذع مع الرّأس لشخصية شهيرة (وفي بعض العملات المحلية يقابلها "الكتابة")، وأمّا على الوجه الآخر فكان ينقش عليه شعار البلد Tail ويماثلها "الشّعار" في العملة المحليّة أيضاً.

٦- لو قمنا بتحليل ضوء الشمس باستخدام موشور، فعندئذٍ يكون لمجموعة الألوان المرئية (ولتكن E مثلاً) التي تنتج عن هذه العملية العرض الآتي:

$E =$ {بنفسجي, أزرق, أخضر, أصفر, برتقالي, أحمر}

٧- إنَّ المقاولين في المملكة العربيّة السعوديّة يكوِّنون مجموعة (ولتكن F مثلاً)، حيث لدينا أنواع كثيرة من المقاولات (مقاولات الطرق، مقاولات البناء، مقاولات التعليم، ...) ولكنّهم جميعاً يتصفون بصفة المقاول، ولذلك يمكننا تمثيلهم في مجموعة على النّحو الآتي:

$F = \{x \mid x$ مقاول $\}$

٨- بالعودة إلى المثال السّابق (١) (رمي حجر نرد سداسي الوجوه لمرّة واحدة فقط) سنقوم بتعيين المجموعات التي تُعبِّر عن حصولنا على:

أ- عدد أكبر من 2.

ب- عدد أكبر أو يساوي 3 وأصغر من 6.

ج- العدد 3.

د- عدد أصغر من 1.

هـ- عدد أكبر أو يساوي 1.

الحلول: لدينا ممّا سبق $A = \{1,2,3,4,5,6\}$، ومن ثمَّ من أجل الطلب:

أ- لنرمز بـ B للمجموعة التي تُعبِّر عن حصولنا على عدد أكبر من 2، فيكون لدينا:

$B = \{3,4,5,6\}$

فنلاحظ أنَّ $B \subset A$ (لاحظ أنَّ العكس غير صحيح، وذلك لأنَّ علاقة الاحتواء $\subseteq$ أو $\subset$ ليست انعكاسيّة).

ب- لنرمز بـ C للمجموعة التي تُعبِّر عن حصولنا على عدد أكبر أو يساوي 3 وأصغر من 6، فيكون لدينا $C=\{3,4,5\}$، فنلاحظ أنَّ $B\supset C$، ومن ثمَّ $A\supset C$ (وذلك لأنَّ علاقة الاحتواء $\supseteq$ أو $\supset$ متعديّة).

ج- لنرمز بـ D للمجموعة التي تُعبِّر عن حصولنا على العدد 3، فيكون لدينا $D=\{3\}$، وهنا نلاحظ أنَّ $C\supset D$.

د- لنرمز بــ E للمجموعة التي تُعبِّر عن حصولنا على عددٍ أصغر من 1، فيكون لدينا:
$E=\{\ \}=\varnothing$

هـ- لنرمز بــ F للمجموعة التي تُعبِّر عن حصولنا على عددٍ أكبر أو يساوي 1، فيكون لدينا:
$F=\{1,2,3,4,5,6\}=A$

ت-١-٤- ملاحظات

١- لاحظ هنا أنَّ المجموعة A أخذت دور المجموعة الشّاملة Ω.

٢- نلاحظ من الأمثلة السابقة أنَّه أمكننا عرض بعض المجموعات من خلال ذكر جميع عناصرها كما في الأمثلة السابقة ١، ٢، ٣، ٥، ٦ و٨، وبعضها الآخر عُرِضَ من خلال كتابة القاعدة التي تُسْتَنْبط منها العناصر كما في المثالين السّابقين ٤ و٧.

ت-١-٥- العمليات المنطقيّة على المجموعات

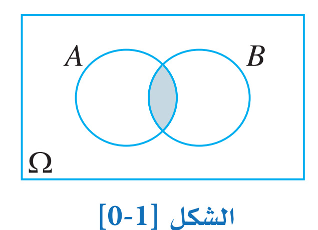
بادئ ذي بدئ نشير إلى بعض الرموز التي يمكن أن تستخدم لدى صياغة المجموعات أو في العلاقات الرّياضياتيّة ومنها مُحدِّدات الكميّة المنطقيّة وهي:

$\forall$ بمعنى "من أجل كلِّ" أو "من أجل أيّ"، ويدعى مُحدِّد الكميّة الكلي (أو الشّامل) Universal Quantifier ، فعلى سبيل المثال $\forall x\in A$ تُقرأ على النّحو الآتي: من أجل أيّ عنصر x من A، أو من أجل كلّ عنصر x من A.

$\exists$ بمعنى "يوجد على الأقل"، ويدعى مُحدِّد الكميّة الوجودي Existential Quantifier، فعلى سبيل المثال $\exists x\in A\ ;\ \cdots$ تُقرأ بالعبارة: يوجد على الأقلّ عنصر x من A بحيث يكون ... .

من المعلوم أنَّه في علم المنطق يمكن تشكيل أي قضية من القضايا الأوليّة باستخدام العمليات المنطقيّة (و- أو- كلا)، وهذه العمليات الثّلاث يقابلها في علم المجموعات العمليات الثّلاث الآتيّة على الترتيب:

١- التَّقاطع Intersection:

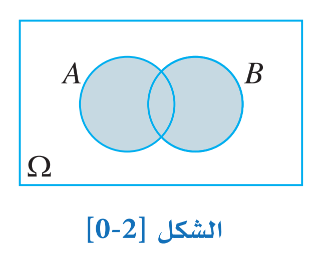


الشكل [0-1]

يُعرَّف **تقاطع** مجموعتين A وB على أنَّه مجموعة العناصر التي تنتمي إلى كلٍ من A وB بآنٍ واحد، ويستخدم الرّمز $\bigcap$ بين المجموعات للدلالة على التّقاطع بينها (أنظر الشكل [0-1])، ومن ثمَّ يُرمز للمجموعة الممثّلة لتقاطع A وB بالرمز $A \bigcap B$، أي أنَّه لدينا:

$$A \bigcap B := \{x | \ x \in A \ \ and \ x \in B \} \qquad [0\text{-}1]$$

٢- الاتحاد Union:

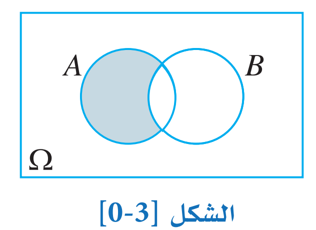


الشكل [0-2]

يُعرَّف **اتحاد** مجموعتين A وB على أنَّه مجموعة كلّ العناصر التي تنتمي إلى A أو إلى B أو إلى كليهما معاً، علماً أنَّ " **أو**" هنا لا تفيد الحصر (أنظر الشكل [0-2])، ويستخدم الرّمز $\bigcup$ بين المجموعات للدلالة على الاتحاد بينها، ومن ثمَّ يرمز للمجموعة الممثّلة لاتحاد A وB بالرمز $A \bigcup B$، أي أنَّه لدينا:

$$A \bigcup B := \{x | \ x \in A \ or \ x \in B \} \qquad [0\text{-}2]$$

٣- الفرق Difference:



الشكل [0-3]

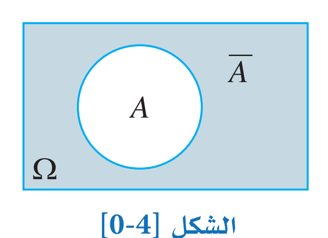
يُعرَّف **الفرق** بين مجموعتين A وB على أنَّه مجموعة العناصر التي تنتمي إلى A ولكن دون B (أنظر الشكل [0-3])، و يُستخدم الرّمز $\setminus$ بين المجموعات للدلالة على الفرق بينها، ومن ثمَّ يرمز للمجموعة الممثّلة لفرق A وB بالرمز $A \setminus B$، أي أنَّه لدينا:

$$A \setminus B := \{ \ x \ | \ x \in A, \ \ x \notin B \ \} \qquad [0\text{-}3]$$

نشير هنا إلى أنَّ كتابة $A - B$ بدلاً من $A \setminus B$ تُعدّ من الكتابات الغامضة وغير المرغوب فيها، وذلك لأنَّ $A - B$ تستخدم عادة للتعبير عن المجموعة الآتيّة:

$$A - B = \left\{ a - b | \ \forall a \in A, \ b \in B \right\}$$

٤- متمّمة مجموعة Complement of a Set:



الشكل [0-4]

لتكن Ω مجموعة شاملة، ولنأخذ $A \subseteq \Omega$ (من الممكن أن تكون $A = \varnothing$). عندئذٍ يُطلق على مجموعة كلّ العناصر من Ω والتي **لا** تنتمي إلى المجموعة A اسم "**متمّمة المجموعة** A" (أو المتمم المطلق للمجموعة A أيضاً)، ويُرمز عـادة لمتمّم المجموعة A بالشكل $\bar{A}$. أي أنَّه لدينا:

$$\bar{A} := \Omega \setminus A \qquad [0\text{-}4]$$

في الحقيقة إنَّ العمليات المنطقيّة على المجموعات (التقاطع - الاتحاد-الفرق)، وبعض الخصائص الأساسية لهذه العمليات (مثل الخاصة التّبديليّة، والتجميعية، والتوزيعية) وبعض العلاقات الشّهيرة المتعلّقة بهذه المواضيع تعدّ الحجر الأساس لفهم علم المجموعات، ولذلك سنقدّمها بشكل موجز ودون الخوض في تفاصيلها.

ت-١-٦- بعض خصائص العمليات الجبريّة على المجموعات

لتكن Ω مجموعة شّاملة، ولنأخذ A، B وC مجموعات جزئيّة من Ω، فعندئذٍ يمكن للقارئ التحقُّق من صحة العلاقات الآتيّة:

١- الخاصة التّبديليّة (أو التّناظريّة) Commutative Property:

إنَّ كلّ من الاتحاد والتّقاطع يتمتَّع بالخاصة التّبديليّة التي يعبرَّ عنها بالعلاقتين الآتيتين:

a) $A \cup B = B \cup A$

b) $A \cap B = B \cap A$

٢- الخاصة التّجميعيّة Associative Property:

إنَّ كلّ من الاتحاد والتّقاطع يتمتَّع بالخاصة التّجميعيّة التي يعبرَّ عنها بالعلاقتين الآتيتين:

a) $(A \cup B) \cup C = A \cup (B \cup C)$

b) $(A \cap B) \cap C = A \cap (B \cap C)$

٣- الخاصة التوزيعية Distributive Property:

إنَّ كلّ من الاتحاد والتّقاطع يقبل التوزيع على الآخر ويعبرَّ عن ذلك بالعلاقتين الآتيتين:

a) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

b) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

٤- قانونا ديمورغان De Morgan's laws:

لدينا العلاقتين الآتيتين محقّقتين دائماً، وتدعيان **قانوني ديمورغان** (نسبة للرّياضياتي الفرنسي أوغستس ديمورغان (1806-1871) Augustus De Morgan):

a) $\overline{A \cup B} = \bar{A} \cap \bar{B}$

b) $\overline{A \cap B} = \bar{A} \cup \bar{B}$

ت-١-٧- ملاحظات

١- يمكن الرجوع إلى الأصل باستخدام متمّم المتمّم حيث لدينا العلاقة الآتيّة محقّقةً دوماً:

$$\bar{\bar{A}} = A$$

٢- لدينا العلاقتين الآتيتين محقّقتين دوماً:

a) $A \setminus B = A \cap \bar{B}$

b) $A \setminus \bar{B} = A \cap B$

◄ ت-١-٨- أمثلة

١- لتكن لدينا المجموعة الشّاملة الآتيّة:

$$\Omega = \{a, b, c, d, e, f, g\}$$

ولنأخذ منها المجموعات الجزئيّة الآتيّة:

$A = \{d, e, f, g\}$ & $B = \{c, d, e, f\}$

$C = \{a, b, e, f, g\}$ & $D = \{a\}$

ولنقم بتعيين المجموعات الآتية:

a) $A \cap B$ & b) $B \cap D$ & c) $A \cup B$

d) $B \cup C$ & e) $C \setminus B$ & f) $\bar{B}$

الحلول: لدينا من أجل الفقرة:

a) $A \cap B = \{d, e, f\}$

b) $B \cap D = \{\} = \varnothing$

c) $A \cup B = \{c, d, e, f, g\}$

d) $B \cup C = \{a, b, c, d, e, f, g\} = \Omega$

e) $C \setminus B = \{a, b, g\}$

f) $\bar{B} = \Omega \setminus B = \{a, b, g\}$

٢- لنقم برمي حجر نرد سداسي الوجوه لمرّةٍ واحدةٍ فقط، فتكون لدينا مجموعة النتائج الآتيّة:

$\Omega = \{1, 2, 3, 4, 5, 6\}$

ولنقم بتعيين المجموعات التي تُعبِّر عن حصولنا على:

أ- عدد أصغر من 3.

ب- عدد أصغر من 3 أو عدد أكبر من 4.

ج- عدد فردي وأصغر من 3.

د- عدد أصغر من 3 وأكبر من 4.

الحلول: من أجل الطّلب:

أ- لنرمز بـ A للمجموعة التي تُعبِّر عن حصولنا على عدد أصغر من 3، فتكون $A = \{1,2\}$.

ب- لنرمز بـ B للمجموعة التي تُعبِّر عن حصولنا على عدد أكبر من 4، فيكون المطلوب:

$A \cup B = \{1, 2\} \cup \{5,6\} = \{1,2,5,6\}$

ج- لنرمز بـ C للمجموعة التي تُعبِّر عن حصولنا على عدد فردي، فيكون المطلوب:

$C \cap A = \{1,3,5\} \cap \{1,2\} = \{1\}$

د- لنرمز بـ D للمجموعة التي تُعبِّر عن حصولنا على عدد أصغر من 3 وأكبر من 4، فيكون لدينا:

$D = \{\} = \varnothing$

لاحظ هنا أنَّ المجموعة Ω أخذت دور المجموعة الشّاملة.

٣- لنقم برمي حجر نرد سداسي الوجوه لمرّتين متتاليتين فتكون لدينا مجموعة النتائج الآتيّة:

$$\Omega = \begin{Bmatrix} (1,1)\ ,\ (1,2)\ ,\ (1,3)\ ,\ (1,4)\ ,\ (1,5)\ ,\ (1,6), \\ (2,1),\ (2,2),\ (2,3),\ (2,4),\ (2,5),\ (2,6), \\ (3,1),\ (3,2),\ (3,3)\ ,\ (3,4),\ (3,5),\ (3,6), \\ (4,1),\ (4,2),\ (4,3),\ (4,4),\ (4,5),\ (4,6), \\ (5,1),\ (5,2)\ ,\ (5,3),\ (5,4)\ ,\ (5,5),\ (5,6), \\ (6,1),\ (6,2),\ (6,3),\ (6,4),\ (6,5),\ (6,6) \end{Bmatrix}$$

ولنقم بتعيين المجموعات التي تُعبِّر عن حصولنا على:

أ- مجموع أصغر أو يساوي 4.

ب- فرق بين الرقمين أكبر من 3.

ج- عدد فردي في المرّة الأولى، وعدد أصغر أو يساوي 3 في المرّة الثّانيّة.

د- عدد في المرَّة الأولى أكبر من العدد الذي حصلنا عليه في المرَّة الثّانيّة.

الحلول: من أجل الطلب:

أ- لنرمز بـــ A للمجموعة التي تُعبِّر عن حصولنا على مجموعٍ أصغر أو يساوي 4، فيكون لدينا:

$$A = \{(1,1)\ ,\ (1,2)\ ,\ (1,3)\ ,\ (2,1),\ (2,2),(3,1)\}$$

ب- لنرمز بـــ B للمجموعة التي تُعبِّر عن حصولنا على فرقٍ بين المركّبة الأولى والثانية (**المركّبة الأولى ناقص المركّبة الثّانيّة**) أكبر من 3 (**أي أكبر تماماً من 3**)، فيكون لدينا:

$$B = \{(5,1),\ (6,1),\ (6,2)\}$$

ج- لنرمز بـــ C للمجموعة التي تُعبِّر عن حصولنا على عددٍ فرديٍّ في المرّة الأولى، وعلى عدد أصغر أو يساوي 3 في المرّة الثّانيّة، فيكون لدينا:

$$C = \{(1,1)\ ,\ (1,2)\ ,\ (1,3),(3,1),\ (3,2),\ (3,3),\ (5,1),\ (5,2)\ ,\ (5,3)\}$$

د- لنرمز بـــ D للمجموعة التي تُعبِّر عن حصولنا على عددٍ في المرَّة الأولى أكبر من العدد الذي حصلنا عليه في المرّة الثّانيّة، فيكون لدينا:

$$D = \begin{Bmatrix} (2,1),\ (3,1),\ (3,2),\ (4,1),\ (4,2),\ (4,3),\ (5,1),\ (5,2), \\ (5,3),(5,4)\ ,\ (6,1),\ (6,2),\ (6,3),\ (6,4),\ (6,5) \end{Bmatrix}$$

إذن، فالمجموعة Ω هنا لعبت دور المجموعة الشّاملة.

ت-١-٩- تعريف (قدرة مجموعة) (Cardinality of a set)

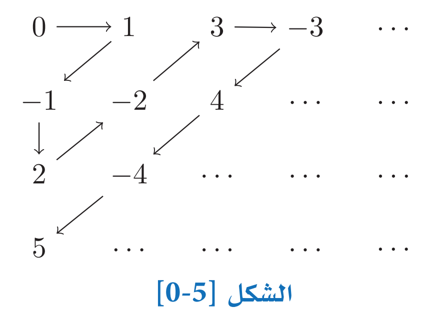
لتكن A مجموعة ما. عندئذٍ تُعرَّف **قدرة** هذه المجموعة بأنَّه العدد الذي يمثّل عدد عناصرها، ويرمز له بـ $|A|$.

فعلى سبيل المثال يكون للمجموعة $A=\{3,-1,7,15,4,-13,2,0,-5\}$ قدرة تساوي 9، وللمجموعة $A=\{a,b,c,d\}$ قدرة تساوي 4.

من المعلوم أنَّ مجموعة ما A قد تكون منتهية (أي تحوي عدداً منتهياً من العناصر) أو غير منتهية، فإذا كان لدينا على سبيل المثال $A=\{a_1, a_2,..., a_n\}$ مع n عدد طبيعي مثبَّت، فإنَّه سيكون لدينا $|A|=n$، وأمّا إذا كانت A غير منتهية فعندئذٍ سنميِّز بين حالتين:

الأولى: إذا كانت المجموعة A غير منتهية ولكن يمكن ترقيم عناصرها بوساطة الأعداد الطّبيعيّة، فعندئذٍ يُقال إنَّ هذه المجموعة **قابلة للعدّ**، ويُكتب $|A|=+\infty$.

لاحظ هنا أنَّ مجموعة الأعداد الطبيعيّة $\mathbb{N}$ هي المعيَّار لقابلية العدّ، فعلى سبيل المثال يمكننا ترقيم عناصر مجموعة الأعداد الصحيحة بوساطة الأعداد الطبيعيّة وفقاً لاتجاه الأسهم في العرض الآتي:



الشكل [5-0]

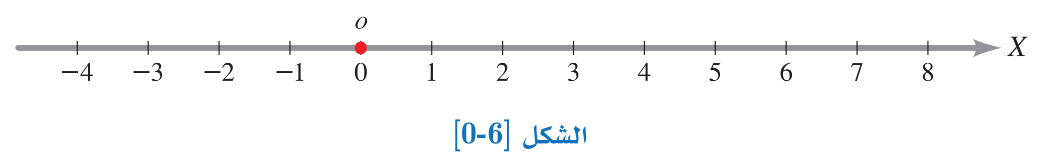
وهذا يعني أنَّ مجموعة الأعداد الصحيحة قابلة للعدّ، وكذلك يمكن إثبات أنَّ مجموعة الأعداد النسبيّة قابلة للعدّ أيضاً.

الثّانيَة: إذا كانت المجموعة A غير منتهية ولا يمكن ترقيم عناصرها بوساطة الأعداد الطّبيعيّة، فعندئذٍ يُقال إنَّ هذه المجموعة **غير قابلة للعدّ**، ويُكتب $|A|=+\infty$ أيضاً، ولكن يُقال في هذه الحالة إنَّ للمجموعة A **قدرة المستمر**.

بالطبع توجد مجموعة تعدّ المعيّار لعدم قابلية العدّ ألا وهي مجموعة الأعداد الحقيقيّة، ولكن التقنية لاستخدام هذا المعيّار لا يمكن إدراجها في هذا الموضع. بمعنى آخر إنَّ مجموعة الأعداد الحقيقيّة هي مجموعة غير قابلة للعدّ، علماً أنَّ مجموعة الأعداد الحقيقيّة هي اتحاد مجموعة الأعداد النسبيّة $\mathbb{Q}$ مع مجموعة الأعداد غير النسبيّة $\mathbb{I}$ ويُرمز لها بـ $\mathbb{R}$، أي أنَّ $\mathbb{R}=\mathbb{Q}\cup\mathbb{I}$.

كما يمكن إثبات أنَّ كلّ فترة I من $\mathbb{R}$ ذات طول موجب تماماً هي مجموعة غير قابلة للعدّ أيضاً.

في الواقع يمكن البرهنة على أنَّه من أجل أي عددين حقيقيين a وb مع $b > a$ يوجد عدد حقيقي c بحيث يكون $b > c > a$، وهذا يعني أنَّ مجموعة الأعداد الحقيقيّة $\mathbb{R}$ **مستمرّة (أو متّصلة)** ولذلك سميّت قدرتها بـ **قدرة المستمر**، وهذه النتيجة هي المبرِّر الحقيقي من أجل التمثيل الهندسي لمجموعة الأعداد الحقيقيّة بالمحور oX حيث يمكن تمثيل الأعداد الحقيقيّة تصويرياً من خلال خط مستقيم موجَّه oX (**أي على شكل سهم**) يكون فيه الرّأس دالاً على تزايد القيم المتوضِّعة عليه، ويُؤخذ (**غالباً**) أفقياً على الشكل الآتي.



الشكل [0-6]

من جهةٍ أخرى يمكن إثبات أن كلّ عدد حقيقي يمكن تخصيصه بنقطة وحيدة على المحور الحقيقي oX (**حيث حرف الـ o هنا يعني نقطة الأصل Origin Point وموضعها ينطبق على الصفر**)، وبالعكس فإذا كان a هو العدد الموافق لنقطة M فإنَّ a تُدعى **إحداثي النقطة** M، ويُشار عادة إلى ترابطهما بالرمز $M(a)$، فعلى سبيل المثال يُشار إلى النقطة M التي إحداثيُّها العدد 3 بالرمز $M(3)$، ولهذا فإنَّنا كثيراً ما نتكلم عن الأعداد بلفظ النقاط، فنتكلم عن النقطة 5 بمعنى النقطة التي إحداثيُّها العدد 5. في هذه الحالة يُدعى الخط المستقيم oX بـــ **خط الإحداثيات** أو **المحور الحقيقي** أو **المستقيم الحقيقي**.

ت-١-١٠- ملاحظات

١- إذا حُذفت مجموعة منتهية من مجموعة قابلة للعدّ فإنَّ الباقي هي مجموعة قابلة للعدّ أيضاً.

٢- إذا حُذفت مجموعة منتهية من مجموعة غير قابلة للعدّ فإنَّ الباقي هي مجموعة غير قابلة للعدّ أيضاً.

٣- إذا حُذفت مجموعة قابلة للعد من مجموعة قابلة للعدّ، فإنَّ الباقي قد يكون مجموعة منتهية أو مجموعة قابلة للعدّ أيضاً.

٤- إذا حُذفت مجموعة قابلة للعد من مجموعة غير قابلة للعدّ فإنَّ الباقي هي مجموعة غير قابلة للعدّ أيضاً.

ت ٢
الدوال الحقيقية

من المفاهيم الهامّة جداً في دراسة الرياضيات ما يُعرف باسم "الدالّة"، حيث يأتي هذا المفهوم في المرتبة الثّالثة من حيث الأهمية بعد مفهوم المجموعة. إنَّ التقديم الآتي يمهّد لنا الطريق للتعرّف على هذا المفهوم.

لنأخذ A وB مجموعتين غير خاليتين، ولتكن f علاقة تُقرن كلّ عنصر a من A بعنصر b من B يُدعى **قيمة** (**أو صورة**) a وفقاً لـ f، ويُرمز لذلك بالشكل $a f b$ أو بأحد العرضين الآتيين أيضاً:

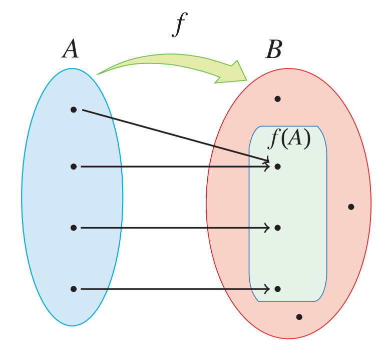
$$f : A \longrightarrow B \;;\; a \mapsto b = f(a) \quad \text{أو} \quad A \xrightarrow[a \mapsto b=f(a)]{} B$$

عندئذٍ يُطلق على هذه العلاقة اسم "**تطبيق** Map" **من** المجموعة A **في** المجموعة B. أمَّا المجموعة A فإنَّها تُدعى **مجال** Domain التطبيق f، في حين يُطلق على المجموعة B اسم "**المجال المقابل** Codomain" للتطبيق f (**وتوجد تسميات أخرى لهاتين المجموعتين في بعض المراجع العربية مثل: منطلق أو مجموعة تعريف أو نطاق بدلاً من كلمة مجال، واستخدام كلمة المُستقر أو النطاق المرافق أو النطاق المصاحب بدلاً من كلمة المجال المقابل**)، وإذا وضعنا:

$$f(A) := \{ b \in B \mid b = f(a); \forall a \in A \} \subset B \qquad [0\text{-}5]$$

فعندئذٍ يُطلق على المجموعة $f(A)$ اسم "**مدى** Range التطبيق f" أو "**صورة** Image التطبيق f " ونرمز لمدى التطبيق f بـ$\mathbf{R}(f)$ (**انظر الشكل** [0-7])، أي أنَّ:

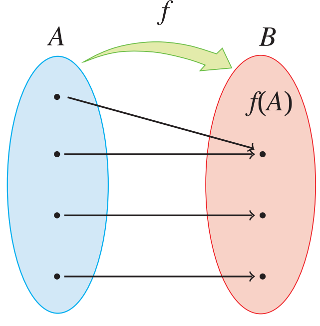
$$\mathbf{R}(f) := f(A)$$



الشكل [0-7] f تطبيق **من** A **في** B.

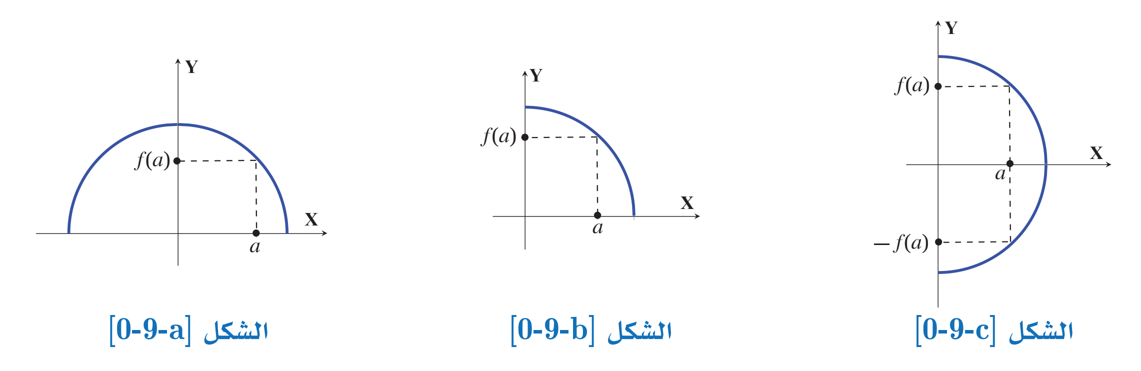
ت-٢-١- ملاحظات

١- في الحالة الخاصّة إذا كانت $f(A)=B$ فعندئذٍ يُقال إنَّ f تطبيق **من** المجموعة A **على** المجموعة B (**انظر الشكل** [0-8])، وفي هذه الحالة يكون المدى $\mathrm{R}(f)$ مساوياً للمجال المقابل.



الشكل [0-8] f تطبيق **من** A **على** B

٢- إنَّ مجموعة كلّ النقاط $\big(a\ ,\ f(a)\big)$ وعندما a تمسح كلّ القيم الممكنة لها في A تُدعى التمثيل البياني للتطبيق f ، فعلى سبيل المثال نجد أنَّ العروض الآتيّة هي تمثيلات بيانية لتطبيقات:



الشكل [0-9-c] الشكل [0-9-b] الشكل [0-9-a]

٣- يُقال عن تطبيق f إنَّه **وحيد القيمة** Single Value إذا كان كلّ عنصر من مجاله يرتبط بعنصر وحيد من مجاله المقابل.

ت-٢-١- تعريف (الدالَّة الحقيقيّة) (Real Function

لتكن A مجموعة ما غير خالية، ولنأخذ:

$$f : A \longrightarrow \mathbb{R}\ ;\ a \mapsto b = f(a)$$

تطبيقاً وحيد القيمة. عندئذٍ يُدعى هذا التطبيق **دالَّة حقيقيّة**.

◄ ت-٢-٢- أمثلة

في العروض البيانية السابقة نجد أنَّ الشكلين [0-9-a] و[0-9-b] يقدِّمان تمثيلين بيانيين لدالَّتين في حين أنَّ الشكل [0-9-c] هو تمثيل بياني لتطبيق ولكنَّه ليس تمثيلاً بيانياً لدالَّة.

►

ت-٢-٣- ملاحظات

١-في معظم الحالات تُستخدم الأحرف اللاتينية كرموز للدوال فيُكتب على سبيل المثال: $f, g , h,...$ أو $F , G , H,...$، وكذلك يُرمز لعناصر مجال الدالّة لهذه الدوال بأحرف لاتينية صغيرة من قبيل $x, y, z,\cdots$ وتُدعى **متغيِّرات مستقلَّة**، وأمَّا القيمة من B والموافقة لـ $f(x)$ فإنَّها تُسمَّى **قيمة الدالَّة** f عند القيمة (**أو النقطة**) x ، ومن المهمّ جداً هنا التمييز بين الدالَّة f التي تمثِّل علاقة وقيمتها $f(x)$ التي تمثِّل عدداً.

٢-إذا كانت f دالَّة معرَّفة على مجموعة $A \subseteq \mathbb{R}$ وتأخذ قيمها في مجموعة $B \subseteq \mathbb{R}$، فعندئذٍ يُقال عن الدالَّة f إنَّها:

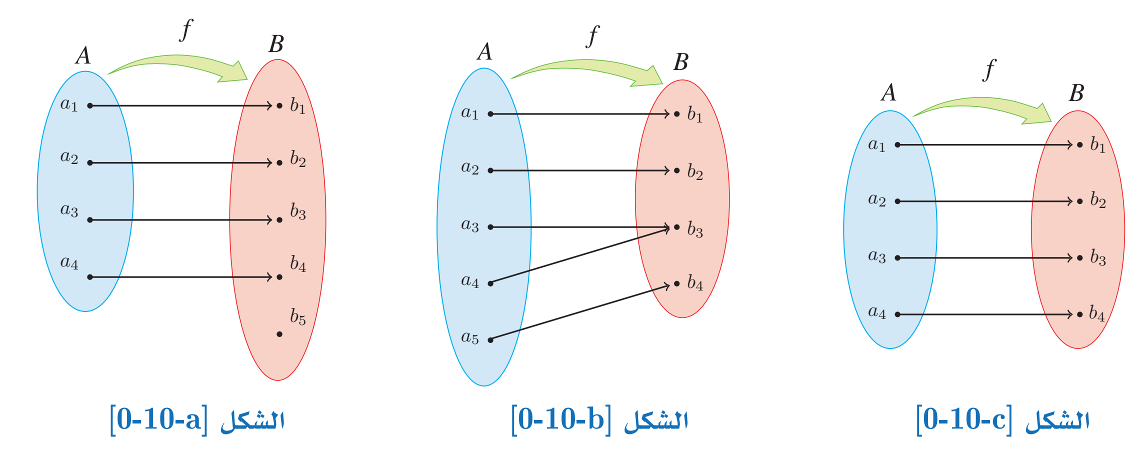
- **متباينة** Injective أو **دالّة واحد لواحدٍ** one-to-one Function (**انظر الشكل [0-10-a]**) إذا كانت هذه الدالّة مُحقّقةً للقضية الآتيّة:

$$f(a)=f(b) \quad \Rightarrow a=b \quad ;\forall a,b \in A$$

- **غامرة** Surjective (**انظر الشكل [0-10-b]**) إذا كانت هذه الدالّة مُحقّقةً للقضيّة الآتيّة:

$$\forall b\in B \ \ \exists\ a\in A \ \ | \ \ b=f(a)$$

- **تقابل** Bijective (**انظر الشكل [0-10-c]**) **إذا وفقط إذا** كانت هذه الدالّة متباينةً وغامرةً في آنٍ واحد.

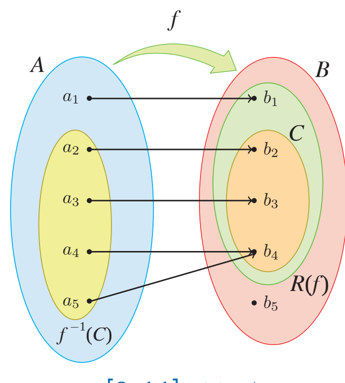


الشكل [0-10-a] الشكل [0-10-b] الشكل [0-10-c]

٣- إذا كانت f دالّة معرّفة على مجموعة $A \subseteq \mathbb{R}$ وتأخذ قيمها في مجموعة $B \subseteq \mathbb{R}$ (أي أنّ f **دالّة حقيقيّة معرّفة على** A)، فعندئذٍ تُعرَّف الصورة العكسيّة لمجموعة $C \subseteq B$ وفقاً للدالّة f على أنّها مجموعة كلّ العناصر a من A التي من أجلها تكون $f(a)\in C$، ويرمز لها عادة بـ $f^{-1}(C)$، أي أنّه لدينا:

$$f^{-1}(C):=\{a\in A \mid f(a)\in C\} \qquad [0\text{-}6]$$

وهي مجموعة جزئيّة من A، فعلى سبيل المثال لو أخذنا الدالّة f المقدّمة في الشكل [0-11] .



الشكل [0-11]

فعندئذٍ نجد أنّ مدى هذه الدالّة هو $R(f)=\{b_1,b_2,b_3,b_4\}$، ولدينا:

$$f^{-1}(\{b_1\})=\{a_1\} \quad \& \quad f^{-1}(\{b_2\})=\{a_2\}$$
$$f^{-1}(\{b_3\})=\{a_3\} \quad \& \quad f^{-1}(\{b_4\})=\{a_4,a_5\}$$
$$f^{-1}(C)=\{a_2,a_3,a_4,a_5\}$$

لاحظ أنَّ الصورة العكسيّة لمدى أيّة دالَّة هي مجال الدالّة كاملاً حيث لدينا من أجل المثال الأخير ما يلي:

$$f^{-1}(f(A))=\{a_1,a_2,a_3,a_4,a_5\}=A$$

ت-٢-٤- تعريف (الدالَّة الخطيّة) (Linear Function)

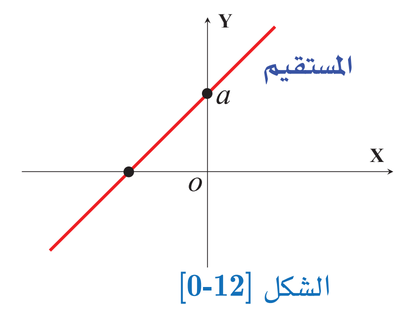
لتكن $A \subseteq \mathbb{R}$، ولنأخذ f دالَّة حقيقيّة معرَّفة على A بحيث تُقرن كلّ عنصر x من A بعدد حقيقي y مُعطى من خلال العلاقة الآتيّة:

$$y = f(x) := a + b\,x \qquad [0\text{-}7]$$

علماً أنَّ a وb ثابتان حقيقيان. عندئذٍ يُقال عن هذه الـدالَّة إنَّها **دالّة خطيّة** على المجموعة A.

ت-٢-٥- ملاحظات

١- إنَّ التمثيل البياني للدالَّة الخطيّة هو مستقيم، ويُقال حينئذٍ عن b إنّه ميل المستقيم (ويساوي ظل الزاوية التي يصنعها المستقيم مع محور الإحداثيّات oX)، في حين تمثِّل النقطة a تقاطع هذا المستقيم مع محور القيم oY (انظر الشكل [0-12]).



الشكل [0-12]

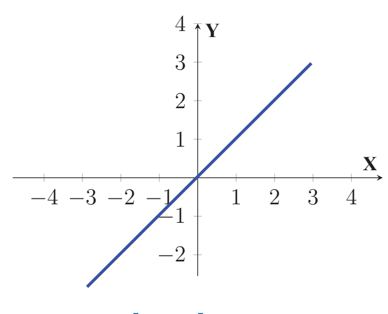
٢- في الحالة الخاصّة عندما تكون قيمة الثابت $a = 0$ فإنَّه يصبح للعلاقة [0-7] العرض الآتي:

$$y = f(x) = b\,x$$

حيث نلاحظ في هذه الحالة أنَّ المستقيم الممثِّل لهذه المعادلة يمرّ من نقطة الأصل (أو من مبدأ الإحداثيات) o، فعلى سبيل المثال لو اخذنا f دالَّة حقيقيّة معرَّفة على $\mathbb{R}$ من خلال العلاقة الآتيّة:

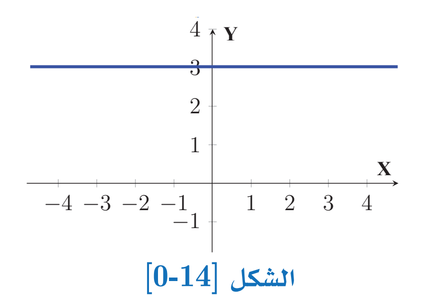
$$y = f(x) = x$$

فعندئذٍ نجد أنَّ ميل المستقيم الممثِّل لهذه المعادلة هو $b = 1$، وأمّا تقاطعه مع محور القيم oY فإنَّه يوافق $a = 0$، ومن ثمَّ يكون للمستقيم الممثِّل لهذه المعادلة الشّكل [0-13] الآتي:



الشكل [0-13]

٣- إذا كان $b = 0$ فإنَّه يصبح للعلاقة [0-7] الصيغة $y = f(x) = a$، وفي هذه الحالة يصبح المستقيم الممثِّل لهذه المعادلة موازياً للمحور الإحداثي **oX** ومارّاً بالنقطة a على محور القيم **oY**، وحينئذٍ يطلق على هذه الدالَّة اسم "**الدالّة الثابتة** Constant Function"، فعلى سبيل المثال لو أخذنا f دالَّة حقيقيّة معرَّفة على $\mathbb{R}$ من خلال العلاقة $f(x) = 3$، فعندئذٍ نجد أنَّ ميل المستقيم الممثِّل لهذه المعادلة هو $b = 0$، وأمّا تقاطعه مع محور القيم **oY** فإنَّه يساوي 3، ومن ثمَّ يكون للمستقيم الممثِّل لهذه المعادلة الشّكل [0-14] الآتي:



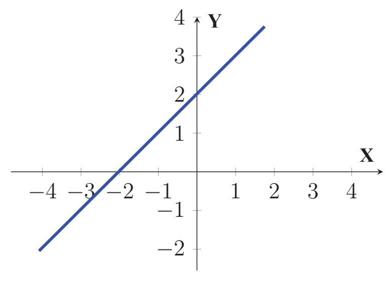
الشكل [0-14]

ت-٢-٦- أمثلة

١- لتكن f دالَّة حقيقيّة معرَّفة على $\mathbb{R}$ من خلال العلاقة الآتيّة:

$$y = f(x) = 2 + x$$

فعندئذٍ نجد أنَّ ميل المستقيم الممثِّل لهذه المعادلة يساوي 1، وأمّا تقاطعه مع محور القيم **oY** فإنَّه يوافق $a = 2$، ومن ثمَّ يكون للمستقيم الممثِّل لهذه المعادلة الشّكل [0-15] الآتي:

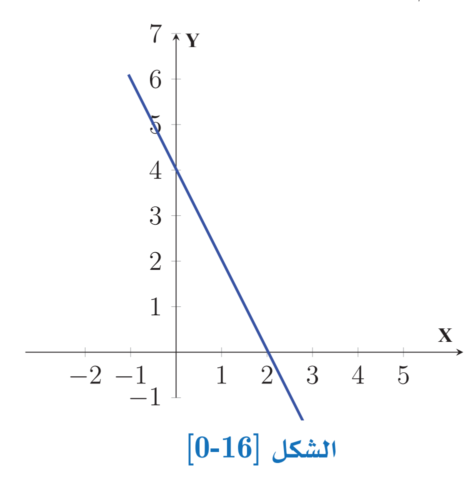


الشكل [0-15]

٢- لتكن f دالَّة حقيقيّة معرَّفة على $\mathbb{R}$ من خلال العلاقة الآتيّة:

$$y = f(x) = 4 - 2x$$

فعندئذٍ نجد أنَّ ميل المستقيم الممثِّل لهذه المعادلة سالب ويساوي -2، وأمّا تقاطعه مع محور القيم **oY** فإنَّه يساوي 4، ومن ثمَّ يكون للمستقيم الممثِّل لهذه المعادلة الشّكل [0-16] الآتي:



الشكل [0-16]

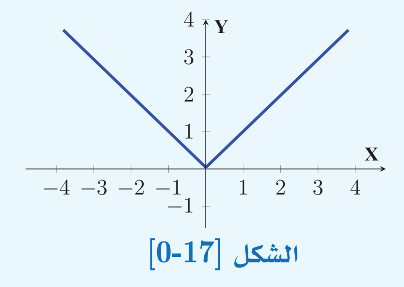
من المثالين السّابقين يُلاحظ أنَّه عندما تكون قيمة الميل موجبة في المعادلة الخطيّة فإنَّه سيكون للمستقيم الممثِّل للمعادلة الخطيّة اتجاه صاعد لدى تزايد قيم x. وعلى العكس، فإذا كانت قيمة الميل سالبة في المعادلة الخطيّة فإنَّه سيكون للمستقيم الممثِّل للمعادلة الخطيّة اتجاه هابط لدى تزايد قيم x.

ت-٤-٧- تعريف (دالَّة القيمة المطلقة Absolut Value Function)

لتكن $A \subseteq \mathbb{R}$، ولنأخذ f دالَّة حقيقيّة تُقرن كلّ عدد حقيقي x من A بعدد حقيقي غير سالب يُرمز له بـ $|x|$، ومُعرَّف على النّحو الآتي:

$$f: A \longrightarrow \mathbb{R} \quad ; x \mapsto f(x) = |x| := \begin{cases} +x & for\ x \geq 0 \\ -x & for\ x < 0 \end{cases} \qquad [0\text{-}8]$$

إنَّ هذه الدالَّة تُدعى **دالَّة القيمة المطلقة**، والعرض البياني لهذه الدالَّة يقدّمه الشكل [0-17] الآتي:



الشكل [0-17]

ت-٢-٨- ملاحظة

ليكن x وy عددين حقيقيين، فعندئذٍ يمكن للقارئ التّحقُّق من صحة العلاقات الآتيّة:

a) $|x| \geq 0$ & $|x| = 0 \Leftrightarrow x = 0$

b) $|x| = |-x|$

c) $|x + y| \leq |x| + |y|$

d) $|x - y| \geq |x| - |y|$

e) $|x \cdot y| = |x| \cdot |y|$

◄ ت-٢-٩- أمثلة

لدينا:

a) $|7| = 7 \quad \& \quad |-5| = 5$

b) $\left.\begin{array}{l} |3+5| = |8| = 8 \\ |3| + |5| = 3 + 5 = 8 \end{array}\right\} \Rightarrow |3+5| = |3| + |5|$

c) $\left.\begin{array}{l} |3+(-5)| = |3-5| = |-2| = 2 \\ |3| + |-5| = 3 + 5 = 8 \end{array}\right\} \Rightarrow |3+(-5)| < |3| + |-5|$

d) $\left.\begin{array}{l} |7-3| = |4| = 4 \\ |7| - |3| = 7 - 3 \end{array}\right\} \Rightarrow |7-3| = |7| - |3|$

e) $\left.\begin{array}{l} |3-7| = |-4| = 4 \\ |3| - |7| = 3 - 7 = -4 \end{array}\right\} \Rightarrow |3-7| > |3| - |7|$

f) $\left.\begin{array}{l} |3 \times 7| = |21| = 21 \\ |3| \cdot |7| = 3 \times 7 = 21 \end{array}\right\} \Rightarrow |3 \times 7| = |3| \cdot |7|$

g) $\left.\begin{array}{l} |3 \times (-7)| = |-21| = 21 \\ |3| \cdot |-7| = 3 \times 7 = 21 \end{array}\right\} \Rightarrow |3 \times (-7)| = |3| \cdot |-7|$

►



تمارين



١- اكتب كلاً من المجموعات الآتيّة بطريقة القاعدة:

- مجموعة أشجار نخيل التمر في المملكة العربيّة السعوديّة.
- مجموعة السيّارات الشّاحنة في مدينة الرياض.
- مجموعة الأعداد الطبيعيّة التي أكبر من 100.
- مجموعة الأعداد النسبيّة بين الصفر والواحد.
- مجموعة الأعداد الأوليّة التي أكبر من 1000.

٢- اكتب كلاً من المجموعات الآتيّة بطريقة العرض السّردي:

- مجموعة كليات العلميّة في جامعة الملك سعود.
- مجموعة البنوك في المملكة العربيّة السعوديّة.
- مجموعة دول مجلس التعاون الخليجي.
- الأعداد الأوليّة التي أصغر من 30.
- مجموعة الأعداد الطبيعيّة التي أصغر من 10.

٣- لتكن $\Omega = \{2,3,5,7,11,13,17,19,23,29\}$ مجموعة شاملة، ولنأخذ المجموعات الجزئيّة الآتيّة من Ω:

$A = \{2,3,5,7,11,13,17\}$

$B = \{7,11,13,17,19,23,29\}$

$C = \{5,7,11,13,17,19\}$

والمطلوب التّحقُّق من صحة العلاقات الآتيّة:

a) $A \cup B = B \cup A$

b) $A \cap B = B \cap A$

c) $(A \cup B) \cup C = A \cup (B \cup C)$

d) $(A \cap B) \cap C = A \cap (B \cap C)$

e) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

f) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

g) $A \cap (B \setminus C) = (A \cap B) \setminus (A \cap C) = (A \cap B) \setminus C$

h) $\overline{A \cup B} = \bar{A} \cap \bar{B}$

i) $\overline{A \cap B} = \bar{A} \cup \bar{B}$

j) $A \setminus B = A \cap \bar{B}$

k) $A \setminus \bar{B} = A \cap B$

l) $(A \setminus B) \cup (B \setminus A) = (A \cap \bar{B}) \cup (B \cap \bar{A})$

٤- لتكن $\Omega = \{a,b,c,d,e,f,g,h,i\}$ مجموعة شاملة ولنأخذ:

$A = \{c, e, g\}$

$B = \{a,c,d,e,f,g\}$

فعندئذٍ تحقَّق من صحة العلاقات الآتيّة:

a) $\bar{B} \subseteq \bar{A}$　　b) $B \setminus A = \bar{A} \setminus \bar{B}$　　c) $A \cap B = A$

d) $A \cup B = B$　　e) $\bar{\bar{A}} = A$

٥- ليكن x عدد حقيقي كيفي، وc ثابت حقيقي موجب، فعندئذٍ تحقَّق من صحة العلاقات الآتيّة مع تمثيل ذلك على محور الأعداد الحقيقيّة:

a) $|x| \leq c \iff -c \leq x \leq c$

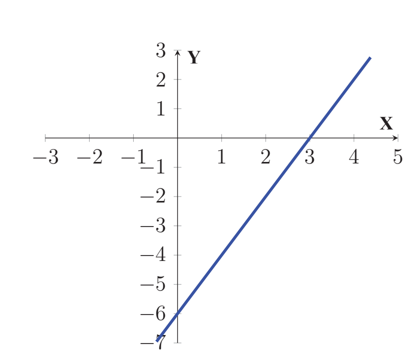
b) $|x| < c \iff -c < x < c$

c) $|x| \geq c \iff x \leq -c \quad and \quad x \geq c$

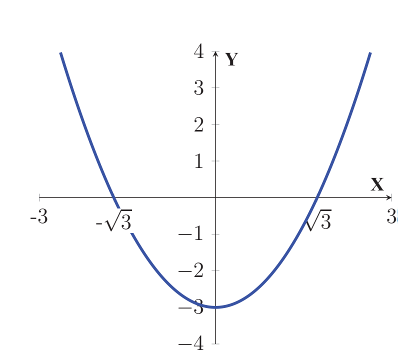
d) $|x| > c \iff x < -c \quad and \quad x > c$

٦- لتكن لدينا العروض البيانية الآتيّة لدوال حقيقيّة:

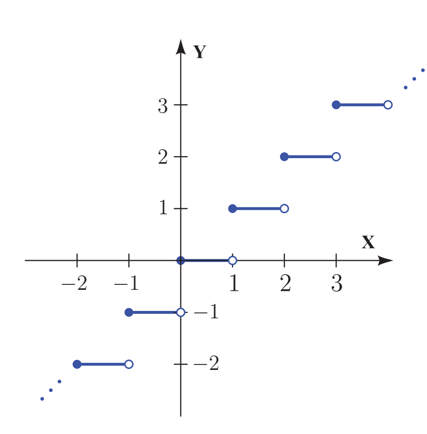
a)



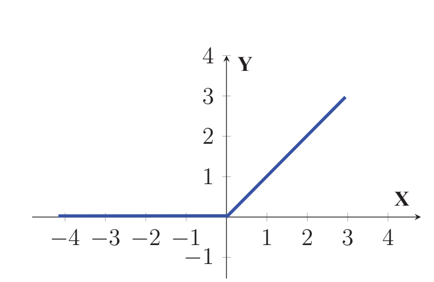
b)



c)



d)



والمطلوب ما يلي:

أ- عيّن المجال والمدى لكلٍّ منها.

ب- وضِّح أي من الدوال المُعطاة لا يمثِّل دالَّة خطيّة.

ج- بيّن أي من الدوال المُعطاة تتمتَّع بخاصّية التقابل، التباين أو الغمور.

الفصل الأول

البيانات الإحصائية جمعها وتنظيمها

Collecting and Organizing Statistical Data



المقدمة:

ينظر إلى علم الإحصاء على أنَّه أحد أهم فروع العلوم الرياضياتية خلال الخمسين سنة الأخيرة، وذلك لأنَّه يلعب دوراً مهمّاً في شتى مجالات الحياة بحيث يكاد لا يخلو أي مجال من مجالات المعرفة من استخدامه. إنَّ الخدمات الكبيرة لعلم العشوائيات دفع علماء الرياضيات قُدماً في تسخير كل قدراتهم لتطوير هذا العلم حتّى يلبي متطلبات التطور التقني المعاصر، ولذا نجد أنَّ الأستاذ الدكتور إيفو شنايدر Ivo Schneider (أستاذ في تاريخ العلوم الطبيعية) يذكر في هذا الصدد (في كتابه تاريخ الاحتمالات منذ البدايات وحتى عام ١٩٣٣)، أنَّه ما من فرع من فروع الرياضيات تطوَّر خلال العقود الأخيرة مثل التطور الذي حظي به فرع العشوائيات (الاحتمالات والإحصاء الرياضي). في الواقع إنَّ علم العشوائيات أوسع بكثير من أن يُلَّم به في كتاب أو مجلد أو حتى عدّة مجلدات، وذلك لكثرة اختصاصاته الفرعية وموضوعاته المتنوّعة.

١ - ١
تعاريف ومفاهيم أساسية

من المعتاد في دراسة أي علم من العلوم التّطرق أولاً إلى تعريف هذا العلم، وذلك لأنَّ أي علم من العلوم يبدأ بتعريفه الذي يعدّ بمثابة حجر الأساس لذلك العلم. لذلك لا بدَّ لنا أولاً أن نتعرف على ما تعنيه عبارة "**علم الإحصاء**". في الحقيقة يمكن تعريف علم الإحصاء على النحو الآتي.

١-١-١- تعريف (علم الإحصاء Statistics)

إنَّ علم الإحصاء هو ذلك الفرع من الرياضيات الذي يهتمّ بجمع البيانات، وتنظيمها (**في جداول وعروضٍ بيانية مناسبة**)، ودراسة خصائصها، وتحليلها، واستقرائها، وأخيراً اتخاذ القرارات المناسبة بشأنها.

كما هو واضح من تعريف علم الإحصاء فإنَّه يمكن تجزئة هذا العلم في قسمين رئيسيين. الأول منهما يهتمّ بجمع البيانات وتنظيمها ودراسة خصائصها العددية (**ويُدعى الإحصاء الوصفي** Descriptive Statistics)، وأمّا الآخر فإنَّه يهتمّ بتحليل البيانات واستقرائها واتخاذ القرارات المناسبة بشأنها (**ويُدعى الإحصاء الاستدلالي** Inferential Statistics، **وفي مراجع أخرى يذكر باسم الاستدلال الإحصائي** Statistical Inference **أيضاً**)، وبناءً على ذلك يمكننا أن نقدِّم التعاريف الآتية.

١-١-٢- تعريف (الإحصاء الوصفي Descriptive Statistics)

هو ذلك الفرع من الإحصاء الذي يهتمّ بجمع البيانات وتنظيمها في جداول وعروض بيانية مناسبة، ودراسة خصائصها العددية.

في الواقع إنَّ هذه الجزئية من العلم يشترك فيها الاختصاصي وغير الاختصاصي أيضاً، فلو أخذنا عملية جمع البيانات الإحصائية يهتمّ المختصون بالطّرائق والأساليب (**أي بالتقنيَة**) التي ستجمع بها البيانات، في حين يمكن لأشخاص غير مختصين (**ولكن مدربين**) القيام بجمع هذه البيانات، فعلى سبيل المثال لو أخذنا عملية المسح السكاني الشامل في بلد ما، فإنَّ دور المختصين هو تحديد الطرائق والأساليب التي ستجمع بها البيانات، وإعداد الاستبيانات المناسبة للأهداف المطلوبة من هذه العملية. في حين يدرَّب معلّمون من المدارس على تعبئة الاستبيانات وكيفية تحصيل البيانات من الأهالي (**لكيلا يحصل التّضليل الإحصائي**). بعد ذلك يقوم الاختصاصي مرّةً أخرى بتنظيم هذه البيانات في جداول وعروض بيانية مناسبة (**لهدف الدراسة الإحصائية**) من أجل أخذ انطباع أولي عن سلوك هذه البيانات.

فيما سبق وردت عبارة "**بيانات**" لمرَّات عديدة، ولكيلا يكون هنا التباس في استخدمها سنقدَّم تعريفها على النحو الآتي.

١-١-٣- تعريف (البيانات) (Data)

البيانات هي قياسات Measures أو ملحوظات (أو مشاهدات) Observations تهدف إلى غرض معيّن في مجتمع إحصائي (سنأتي على تعريفه بعد قليل) مُحدّد، ويتمّ تدوينها كنتيجة لعمليّة إنتاجيّة (مثل كميّات القمح الناتج عن الزّراعة في عام أو أعوام متتالية في بلد ما) أو لتجربة معمليّة (مثل معرفة الألوان النّاتجة عن تحليل ضوء الشمس) أو لمراقبات (ملاحظات) لكائنات موجودة في الطبيعة (مثل أعداد النجوم في مَجرّة في الفضاء الكوني) أو ....

١-١-٤- تعريف (الإحصاء الاستدلالي) (Inferential Statistics)

هو ذلك الفرع من الإحصاء الذي يهتمّ بتحليل البيانات واستقرائها (تعميم نتائج العينات على المجتمع الإحصائي) ومن ثمَّ اتخاذ القرارات المناسبة بشأنها.

إنَّ العمل في هذا القسم من الإحصاء يقع على عاتق الاختصاصيين حصراً، وذلك لأنَّها تتطلب مهارات علميَّة على مستوى أكاديمي، حيث يبدؤون بتحليل البيانات ودراسة توزيعاتها وتعيين مَعْلماتها ومن ثمَّ البحث عن مقدِّرات مناسبة لهذه المعالم وبعد ذلك دراسة سلوك تلك المقدِّرات و.... وأخيراً اتخاذ القرارات المناسبة بشأن المجتمع قيد الدراسة. بالطبع سوف لن نتناول هذه الدراسة في كتابنا هذا لأنَّها خارج إطار خطته.

في هذا الفصل من كتابنا هذا سوف نتناول أول وأبسط جزئية من علم الإحصاء ألا وهي "البيانات الإحصائية جمعها وتنظيمها" كما ورد في عنوان هذا الفصل، وهذه الدراسة تتطلب العديد من المفاهيم الإحصائية وأولها هو الآتي.

١-١-٥- المجتمعات الإحصائية Populations

في الواقع إنَّ الدّراسة الإحصائية لأية مسألة تنطلق ممّا يُعرف باسم "المجتمع الإحصائي" الذي يكوِّن الركيزة الأساس للبيانات التي ستخضع للدراسة، ولذلك لابدّ من تقديم تعريف واضح لمعنى المجتمع الإحصائي كي لا يكون هناك أي لبس في ذهن القارئ لهذا المفهوم المهم.

١-١-٥-١- تعريف (المجتمع الإحصائي) (Population)

المجتمع الإحصائي هو أي تجمُّع لأشياء تَجمَع بينها صفة مشتركة واحدة على الأقل لتكون محل دراسة لهدف محدَّد.

١-١-٥-٢- ملاحظات

١- سنستخدم كلمة "مجتمع" عوضاً عن "مجتمع إحصائي" على سبيل الاختصار والتبسيط، وإذا ما كُتبت بين الحين والآخر فإنَّ ذلك من باب التذكير بها فقط.

٢- تُدعى مكوّنات المجتمع عناصر أو أفراداً.

٣- إنَّ عدد عناصر المجتمع يُدعى **حجم المجتمع**، ولذلك فمن الممكن أن يكون حجم المجتمع:

- **محدوداً**، وفي هذه الحالة يكون عدد عناصر المجتمع منتهياً ويُرمز له بعدد طبيعي، وقد درجت العادة على استخدام أحرف لاتينية كبيرة من قبيل N ، M و... للدلالة على حجم المجتمع، فعلى سبيل المثال من أجل دراسة إحصائية ما على طلاب جامعة الملك سعود يمكن النظر إلى طلاب هذه الجامعة على أنَّه مجتمع محدود.
- **غير محدود**: وفي هذه الحالة يكون عدد عناصر المجتمع غير منتهٍ، ولذلك لا يستخدم رمزاً للدلالة على حجم المجتمع في هذه الحالة، فعلى سبيل المثال من أجل دراسة إحصائية ما على سلوك الأعداد الأوليّة (مثل عشوائيتها وتوزيعها) فإنَّه يمكن النظر إلى مجموعة الأعداد الأولية على أنَّها مجتمع غير محدود.

٤- من الأمور المهمّة هنا هي أن ندرك أنَّ المجتمع ليس بالضرورة أن يكون مجتمعاً بشرياً أو حتى مجتمعاً لأحياء، إذ إنَّه من الممكن أن يكون جماداً أو أي شيء آخر أيضاً، والمثالان الآتيان يوضّحان لنا ذلك.

- لدى تحديد نسبة خام النّحاس في فلز معدني في منجم معيّن، حيث يمكن أن تتواجد أنواع عديدة من مركّبات النحاس في هذا المنجم، ولكنَّها جميعاً تحوي على معدن النحاس، ولذلك فلز النحاس في هذا المنجم يكوِّن مجتمعاً.
- في دراسة لتحديد الحالة الفنيّة للطائرات السّفريّة في المملكة العربية السعودية، حيث يوجد أنواع عديدة من الطائرات السفرية، ولكنَّها جميعاً تتصف بأنَّها طائرات سفرية، ولذلك الطائرات السفرية الموجودة في المملكة العربية السعودية تكوِّن مجتمعاً.

١-١-٦- العيِّنات Samples

قد تكون عملية إخضاع جميع عناصر المجتمع للبحث والدراسة شاقة، وأكثر من ذلك قد تكون في كثير من الحالات غير ممكنة أيضاً، فعلى سبيل المثال:

- لو أرادت هيئة الرّقابة على الأدوية التحقُّق من مكوّنات عقار دوائي معيّن معبأ في كبسولات، فعندئذٍ من غير المجدي أن تقوم هذه الهيئة بإخضاع كل إنتاج المصنع (ومن ثمَّ إتلاف كافة الإنتاج) للتحليل المخبري من أجل التثبت من أنَّ المنتج مُحققاً للمواصفات المقدَّمة من قِبل المصنع.
- في عملية تحليل الدم لمريض فمن غير المعقول ولا المقبول أخذ كل دم المريض (ومن ثمَّ قتل المريض) لتحليله من أجل الكشف على أسباب مرضه.

بالطبع هناك عوامل أخرى قد تضطرُّنا إلى عدم التعامل مع عناصر المجتمع كلّه لأسباب أخرى منها الاقتصاديّة والزمنيّة أيضاً، ولذلك يلجأ المرء في مثل هذه الحالات إلى أخذ جزء من المجتمع لدراسته، وهذا العمل يندرج تحت مفهوم العينة والذي يقدِّمه لنا التعريف الآتي.

١-١-٦-١- تعريف (العيِّنة) (Sample)

تُعرَّف العيِّنة على أنها جزء من المجتمع يتمّ اختياره بشكل مناسب بحيث يُمثِّل المجتمع تمثيلاً جيّداً.

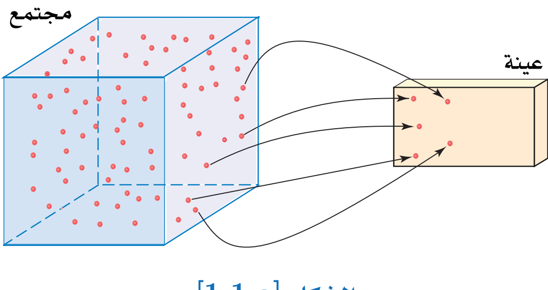
١-١-٦-٢- ملاحظات

١- سنستخدم كلمة عيِّنة عوضاً عن كلمة عيِّنة إحصائيّة على سبيل الاختصار والتبسيط.

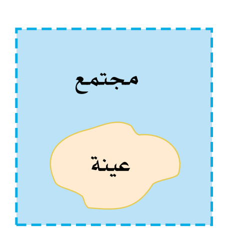
٢- نشير إلى أنَّ عدد عناصر العيِّنة يجب أن يكون منتهياً.

٣- يُطلق على عدد عناصر العيِّنة اسم "**حجم العيِّنة**"، وقد درجت العادة على استخدام أحرف لاتينية صغيرة من قبيل n، m و... للدلالة على حجم العيِّنة.

الشكلان الآتيان [1-1-a] و[1-1-b] يوضّحان مفهوم المجتمع والعيِّنة.



الشكل [1-1-a]



الشكل [1-1-b]

١-١-٦-٣- تصنيف العيِّنات Classification of Samples

مما تقدّم يتبيَّن لنا أنَّ استخدام مفهوم العيِّنة في الدراسة (أو البحث) إنمَّا هو وسيلة لتعميم ما تمَّ التوصل إليه من نتائج على المجتمع، وذلك لاعتقادنا أنَّ هذه العيِّنة هي مُمثِّل مقبول للمجتمع. هذا وتصنَّف العينات في نوعين رئيسيين هما:

١- **عيِّنات عشوائيّة** Random Samples، ويتميَّز هذا النوع من العيِّنات بأنَّ عناصرها تسحب من المجتمع بطرائق عشوائيَّة، ومن أهمّ أنواعها وأكثرها انتشاراً **العيِّنات العشوائية البسيطة**، وهذا النوع من العينات يتميَّز بأنَّ لجميع عناصرها النصيب نفسه في الاختيار (أو السحب أو الانتقاء) مع أي عيِّنة عشوائية أخرى بذات الحجم وممكنة التشكيل من المجتمع نفسه، وهذا يوافق الحالة التي يكون فيها جميع عناصر المجتمع مستقلَّة بعضها عن البعض الآخر ولها النصيب نفسه في الظهور (أو الاختيار).

٢- **عيِّنات عمديَّة** (أو قصديَّة) Intentional Samples ويتميَّز هذا النوع من العيِّنات بأنَّ عناصرها تُنتقى من المجتمع وفقاً لرأي الباحث وخبرته، ولهذا السبب فإنَّه من النادر استخدام هذا النوع من العينات لأنَّه من الممكن أن يتحيّز الباحث في عملية الانتقاء.

١-١-٧- المتغيِّرات Variables

لقد لاحظنا أنَّه لدى أية دراسة إحصائية نكون أمام هدفٍ مُحدَّد المَعالم نأمل الوصول إليه، ومن أجل ذلك كنّا نقوم بتطبيق أداة معيَّنة على أفراد العينة أو المجتمع للحصول على البيانات التي نرغب بها من أجل الوصول إلى قرار بشأن دراستنا الإحصائية. في الواقع إنَّ هذه الأدوات التي ذكرناها تندرج تحت أحد المفاهيم المهمّة في الرياضيات (ألا وهي التطبيقات)، وتُقدَّم في الإحصاء تحت مسمّى المتغيِّرات التي يقدِّمها لنا التعريف الآتي.

١-١-٧-١- تعريف (المتغيِّر Variable):

يُعرَّف المتغيِّر على أنَّه تطبيق (وقد يكون دالّة) مجاله (أو مجموعة قيمه) العيِّنة أو المجتمع نفسه (حسب طبيعة الدراسة الإحصائية)، وأمَّا مجاله المُقابل فهو مجموعة ذات طبيعة ما، فيمكن لها أن تكون أعداداً أو رموزاً أو مسمّيات، ويستخدم لقياس خاصيّة معيّنة لعناصر العيِّنة أو المجتمع.

١-١-٧-٢- ملاحظة:

من التّعريف السابق يتضّح لنا أنَّ القياسات أو المشاهدات التي يمكن أن تنتج عن متغيرٍّ قد تكون قيماً عدديةً أو أحرفاً أو رموزاً أو... وبناءً على ذلك يمكننا تصنيف المتغيِّرات في نوعين رئيسيين هما:

١- **المتغيِّرات الكميَّة** Quantitative Variables، وهي متغيِّرات تكون قيمها أعداداً حقيقية تنتج عن التساؤل بـ " **كم**". إنَّ البيانات التي تنتج عن هذه المتغيِّرات تُدعى **بيانات كميَّة** Quantitative Data، فعلى سبيل المثال:

أ- المتغيِّر الذي يرصد أعداد الطلاب في الجامعات السعودية تكون قيمه أعداداً طبيعيةً ينتج عن التساؤل بـ "كم"، فلو كان عدد طلاب جامعة الملك سعود يساوي 50000 طالب، فعندئذٍ تنتج هذه القيمة عن السؤال: كم عدد طلاب جامعة الملك سعود؟

ب- المتغيِّر الذي يرصد أعداد الأطفال لدى الأسر في مدينة الرياض تكون قيمه أعداداً صحيحة غير سالبة ينتج عن التساؤل بـ "كم"، فلو كان لدى أسرة X من مدينة الرياض خمسة أطفال، فعندئذٍ هذه القيمة تنتج عن السؤال: كم عدد الأطفال لدى الأسرة X؟

ج- المتغيِّر الذي يرصد الطول (مقدَّراً بالسنتيمتر) لدى أفراد بلد ما، فإنَّ قيم هذا المتغيّر هي أعداد حقيقية، ومن الممكن أن يكون لمجاله الفترة [45 , 246]، وينتج عن التساؤل بـ "كم"، فلو أخذ شخص ما X من ذلك البلد وكان طوله 176 سنتيمتر، فعندئذٍ هذه القيمة تنتج عن السؤال: كم طول الشخص X؟

من هذه الأمثلة نلاحظ أنَّه يمكننا تصنيف المتغيِّرات الكميَّة في نوعين أيضاً، وهما:

١-أ- **متغيِّرات متقطِّعة** Discrete Variables، وهي تلك المتغيِّرات الكميَّة التي مجالها (**مجموعة قيمها**) منتهٍ أو غير منتهٍ ولكن قابلة للعدّ، ومن الأمثلة على ذلك:

- المتغيِّر الذي يرصد عدد السّيارات المباعة من معرض ما في يوم معيَّن حيث يكون لمجاله المجموعة $\{0,1,2,3,...,n\}$ مع n عدد طبيعي مثبَّت.
- المتغيِّر الذي يرصد عدد التجارب التي يجب تنفيذها حتى الحصول على شعار لأول مرَّة لدى قذف قطعة نقود معدنيّة، فنجد أنَّ لمجاله المجموعة $\{1,2,3,...\}$ وهي مجموعة الأعداد الطبيعية كاملة.

١-ب- **متغيِّرات مستمرّة (أو متّصلة)** Continuous Variables، وهي تلك المتغيِّرات الكميَّة التي مجالها (**مجموعة قيمها**) غير قابلٍ للعدّ (**وبالتالي غير منتهة أيضاً**)، ومن الأمثلة على ذلك:

- المتغيِّر الذي يرصد عمر الإنسان في القرن الأخير 2016-1916، فنجد أنَّ لمجاله (**لمجموعة قيمه**) الفترة [0 , 179]، وهي مجموعة غير قابلة للعدّ.
- المتغيِّر الذي يرصد الوقت المستّغرق من قبل طالب لإنهاء اختباره (**بزمن 120 دقيقة**) في مقرَّر معيَّن، فنجد أنَّ لمجاله الفترة [120 , 0)، وهي مجموعة غير قابلة للعدّ.
- المتغيِّر الذي يرصد طول الطفل عند الولادة في مستشفى للتوليد، فنجد أنَّ لمجاله الفترة [20 , 45] على وجه التقريب (**أطراف الفترة بالسنتيمتر**)، وهي مجموعة غير قابلة للعدّ.

٢- **المتغيِّرات النوعيَّة** Qualitative Variable، وهي متغيِّرات تكون قيمها عبارة عن رموز أو أسماء أو أرقام دالَّةً على نوع أو اسم أو صفة أو تمييز، وهذه القيم تنتج عن السؤال بـ "**ما**".

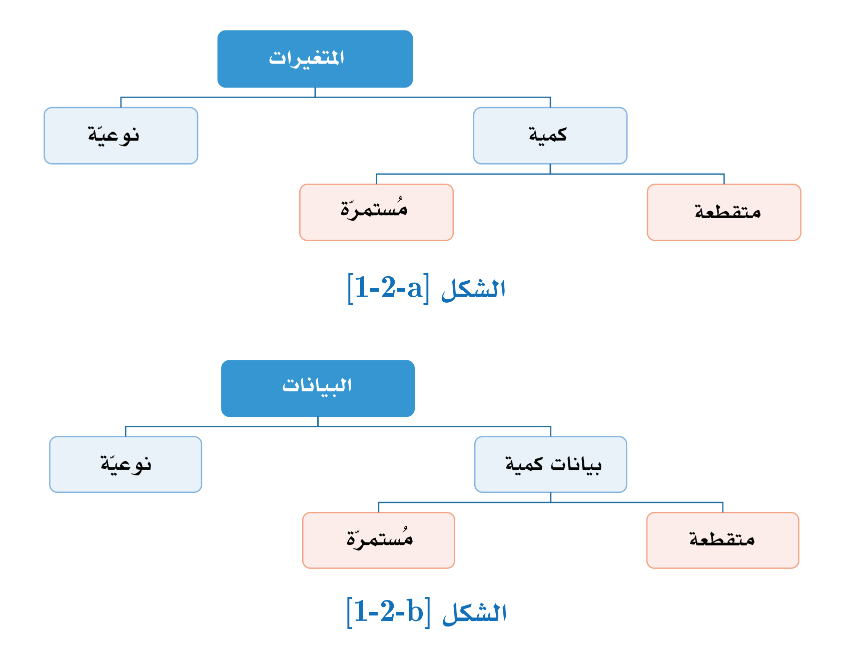
إنَّ البيانات التي تنتج عن هذا النوع من المتغيِّرات تُدعى **بيانات نوعيَّة** Qualitative Data، فعلى سبيل المثال:

أ- المتغيِّر الذي يرصد ألوان الزهور في حديقة معيَّنة هو متغيِّر نوعي، والقيم التي تنتج عنه (**أحمر، أصفر، أبيض و...**)، ونحصل عليها بالسؤال: **ما** لون الزهرة؟

ب- المتغيِّر الذي يرصد فصيلة الدم لدى البشر تكون مجموعة قيمه رموزاً O، A، B وAB، ونحصل على هذه القيم بالسؤال: **ما** فصيلة دم الشخص X؟

ج- المتغيِّر الذي يرصد الرقم الجامعي لطالب في جامعة الملك سعود هو متغيِّر نوعي، والقيم التي تنتج عنه هي أرقام من قبيل ......436، ......437 و.... وهذه الأرقام تميِّز الطّالب ولا تعني مقداراً كميّاً له، ونحصل على هذه القيم بالسؤال: **ما** رقم الطالب X؟

الشّكلان الآتيان يوضّحان لنا أنواع المتغيِّرات والبيانات.



الشكل [1-2-a]

الشكل [1-2-b]

١ - ٢
تنظيم البيانات الخام وتمثيلها

الآن، وبعد إتمام عملية جمع البيانات تأتي المرحلة التّالية، وهي البحث في كيفية التعامل مع هذه البيانات. حيث يتمّ التحقُّق أولاً من اكتمال البيانات، فإذا كانت هناك بعض البيانات المفقودة فعندئذٍ يجب النّظر في كيفية تحصيلها ثانية أو تقديرها إن أمكن ذلك، وإلاّ سوف تُستثنى من الدراسة (علماً أنّ عملية التقدير للبيانات المفقودة تقع خارج إطار هذا الكتاب، ولذلك سوف لن نتطرق إلى عملية تقدير البيانات المفقودة).

إنّ البيانات التي نحصل عليها قد تكون على أشكال مختلفة، فمنها على شكل قيم عددية مُفْردة، وبعضها الآخر قد يَعْرِض تَغيُّرَ ظاهرةٍ ما مع مرور الزمن أو مع مسمَّيات كالبلدان، أو المدن، أو مع كليهما معاً، وبعضها الآخر قد يكون مجمَّعاً في جداول. لذلك سنبحث في تنظيم البيانات وفق اتجاهين:

الأول: يهتمّ بتنظيم البيانات المُفْرَدة التي تنتج مباشرة عن الدراسة الإحصائية (كميَّةً كانت أم نوعيّة) في جداول تُدعى الجداول التكرارية، ومن ثمَّ تمثيل هذه البيانات في عروض بيانية مناسبة.

الثاني: يهتمّ بتجميع البيانات المُفْرَدة الكميَّة فقط في جداول من نوع خاص تُدعى جداول التوزيع التكرارية، حيث يُقال عن البيانات المقدَّمة بهذه الجداول إنَّها بيانات مجمَّعة (أو مبوَّبة، أو مجدولة)، ومن ثمَّ تمثيل بيانات هذه الجداول في عروض بيانية مناسبة.

١-٢-١- البيانات الخام Raw Data

لدى تنفيذ دراسة إحصائية معيَّنة حول ظاهرة ما وجمع البيانات حول هذه الظاهرة تنتج لدينا بيانات مُفْرَدة تُدعى **بيانات خام**.

إذا كان عدد البيانات صغيراً فإنَّه يمكن التعامل مع هذه البيانات بشكل مباشر (مع كل قيمة على انفراد) لدراستها، وأمّا إذا كان عدد البيانات كبيراً نسبياً، فإنَّه قد يكون من الصعب التعامل معها بشكل مباشر، ولذلك لا بدَّ من تقديم طرائق تسهل التعامل مع هذه البيانات كي يتمّ الاستفادة منها بأقصى قدر ممكن.

فيما يلي نقدِّم بعض الأمثلة على بيانات خام قبل الشروع في تقديم طرائق عرضها جدولياً.

١-٢-١-١- أمثلة

١- لدى الاطلاع على تقديرات 60 طالباً في مقرَّر الإحصاء وجدنا المُعطيات الآتية:

C	A	D	B	D	A	F	D	C	A
D	D	D	C	C	B	C	F	F	C
A	B	C	D	D	A	D	A	A	B
D	C	F	D	C	B	C	C	B	C
B	D	A	B	B	C	B	A	D	C
C	C	F	C	B	D	C	D	B	F

فنلاحظ أنَّ هذه البيانات هي بيانات خام نوعيّة.

٢- لقد سُئل 30 طالباً من كلية X عن عدد الحوادث المرورية التي حصلت معه خلال الفصل الدراسي الأول لهذا العام فكانت الإجابات كما يلي:

0	0	1	3	1	0	1	2	2	3
2	0	1	2	1	1	1	1	2	1
1	0	0	0	3	2	2	1	0	0

فنلاحظ أنَّ هذه البيانات هي بيانات خام كميّة.

٣- البيانات الآتية تمثِّل الطول لخمسين طالباً جامعيّاً (مقدَّرة بالسنتيمتر):

140	155	168	171	158	168	159	149	172	145
155	154	166	169	168	158	149	172	168	166
156	166	149	157	156	159	167	166	169	171
170	159	168	168	167	157	154	166	169	168
158	157	172	155	154	166	168	167	171	168

وهذه البيانات هي بيانات خام كميَّة، ولكنَّها تتبع مجموعة بيانات مستمرّة (أو متّصلة). حيث نلاحظ أنَّ كمية البيانات الخام المقدَّمة أعلاه لا تعدّ كبيرة نسبياً إلاَّ أنَّه يصعب أخذ انطباع سريع عن سلوك هذه القيم بشكلٍ مباشر، فعلى سبيل المثال: هل كل قيمة في هذه المجموعة تتكرَّر بالقدر نفسه؟

بالطّبع قد يقوم المرء بإجراء تعداد لكل بيان من هذه البيانات ومن ثمَّ إجراء مقارنة بين هذه التّعدادات، ولكنَّ ذلك سيستغرق وقتاً لا بأس به. لذلك اقترح تنظيم هذه البيانات بطريقة معيَّنة حتى يسهل على المرء الاستفادة منها واستنباط سلوك هذه البيانات بأقل وقت وجهد ممكن. من هنا جاءت فكرة صبّ البيانات في جداول يذكر في أحد أعمدتها البيان المميَّز (الرمز أو العدد الممثِّل)، وفي العمود المقابل يذكر التعداد وفي العمود الذي يليه يدوَّن عدد يعبِّر عن تعداد هذا البيان.

١-٢-٢- التّمثيل الجدولي للبيانات الخام Table Representation of Raw Data

إنَّ تمثيل البيانات الخام جدولياً يعني صبّ هذه البيانات في جدول بتصميم معيَّن، وهذا الجدول يُدعى **الجدول التكراري** للبيانات. فإذا أردنا صبّ مجموعة بيانات خام في جدول تكراري نقوم بإدراج جدول يحتوي على خمسة أعمدة، وهذه الأعمدة تُخصَّص على النحو الآتي:

أ- يدوَّن في العمود الأول **ممثِّل** لكل نوع من الأسماء أو الرموز أو الأعداد حسب طبيعة البيانات التي قيد الدراسة، فعلى سبيل المثال لدينا:

- الرمز A هو ممثِّل لكل الرموز A في مجموعة بيانات المثال (١) من الفقرة (١-٢-١-١)،
- العدد 0 هو ممثِّل لكل الأعداد 0 في مجموعة بيانات المثال (٢) من الفقرة (١-٢-١-١).

ب- يدوَّن في العمود الثّاني **التّعداد** Tally لكل ممثِّل (**من الأسماء أو الرّموز أو الأعداد**) في مجمل البيانات التي قيد الدراسة، ويتمّ ذلك برسم خط عمودي عن كل بيان موافق للاسم أو الرمز أو العدد، وإذا أصبح لدينا أربعة خطوط عمودية فإنَّ الخط الخامس يحزمها على النحو 𝍸.

ج- يدوَّن في العمود الثّالث عدد يُعبِّر عن تعداد كل ممثِّل (**من الأسماء أو الرموز أو الأعداد**) في مجمل البيانات التي قيد الدراسة، وهذا العدد يُدعى **التكرار** Frequency.

د- يدوَّن في العمود الرّابع عدد يعبِّر عن نسبة تكرار كل ممثِّل (**من الأسماء أو الرموز أو الأعداد**) إلى العدد الكلي للبيانات التي قيد الدراسة، وهذا العدد يُدعى **التّكرار النّسبي** Relative Frequency. أي أنَّ التكرار النسبي يساوي تكرار النوع مقسوماً على المجموع الكلي للتكرارات.

هـ- يدوَّن في العمود الخامس عدد يعبِّر عن حاصل ضرب العدد 100 في التكرار النسبي لكل ممثِّل (**من الأسماء أو الرموز أو الأعداد**) ويقرأ كنسبة مئوية، وهذا العدد يُدعى **التكرار المئوي** Percent Frequency. أي أنَّ التكرار المئوي يساوي التكرار النسبي مضروباً في 100.

١-٢-٢-١- أمثلة

١- لنقم بصبّ البيانات الموجودة في المثال (١) من الفقرة (١-٢-١-١) في جدول تكراري، فنجد أنَّ لهذا الجدول العرض الآتي:

الجدول [1-1-a]

التقدير	التعداد	التكرار	التكرار النسبي	التكرار المئوي
A	𝍸 ||||	9	$9/60 = 0.15$	$0.15\times100 = 15\ \%$
B	𝍸 𝍸 ||	12	$12/60 = 0.20$	$0.20\times100 = 20\ \%$
C	𝍸 𝍸 𝍸 |||	18	$18/60 = 0.30$	$0.30\times100 = 30\ \%$
D	𝍸 𝍸 𝍸	15	$15/60 = 0.25$	$0.25\times100 = 25\ \%$
F	𝍸 |	6	$6/60 = 0.10$	$0.10\times100 = 10\ \%$
Total	----------------	**60**	**1**	**100**

نلاحظ هنا سهولة الحصول على المعلومة الخاصّة بالتكرارات، فعلى سبيل المثال إذا أردنا الحصول على عدد الطلاب الحاصلين على تقدير C نجده بكل سهولة يساوي 18، وأنَّ نسبة الطلاب الذين حصلوا على تقدير A يشكلون 15% من إجمالي عدد الطلاب، وأنَّ 25% من الطلاب حصلوا على التقدير D.

٢- لتكن لدينا البيانات الآتية والناتجة عن فحص فصيلة الدم لستين شخصاً.

B	A	B	A	B	O	A	O	AB	A
A	O	A	AB	O	A	AB	O	A	AB
A	B	B	B	A	AB	O	A	AB	A
B	AB	A	A	AB	A	A	O	B	B
AB	A	B	O	A	B	A	AB	A	AB
A	B	A	A	AB	A	O	A	B	B

فلو قمنا بصبِّ هذه البيانات في جدول تكراري على النحو السابق، فإنّنا سنجد له العرض الآتي:

الجدول [1-1-b]

رمز فصيلة الدم	التعداد	التكرار	التكرار النسبي	التكرار المئوي
A	𝍸 𝍸 𝍸 𝍸 ||||	24	$24/60=0.40$	$0.40\times100=40\ \%$
B	𝍸 𝍸 𝍸	15	$15/60=0.25$	$0.25\times100=25\ \%$
AB	𝍸 𝍸 ||	12	$12/60=0.20$	$0.20\times100=20\ \%$
O	𝍸 ||||	9	$9/60=0.15$	$0.15\times100=15\ \%$
Total		**60**	**1**	**100**

٣- في إحدى المدارس أخذت عيّنة مكوَّنةً من 40 طالباً بغية دراسة عدد المرّات التي أصيب فيها الطالب بنزلة برد (انفلونزا) خلال موسم الشّتاء في عام 1438 هـ، فكانت النّتائج كما يلي:

1	0	1	3	2	1	3	2	1	0
2	2	0	1	0	1	2	0	2	1
1	2	1	3	1	2	1	2	1	1
1	0	1	1	2	0	0	1	1	2

فلو قمنا بصبِّ هذه البيانات في جدول تكراري، فإنّنا سنجد له العرض الآتي.

الجدول [1-2]

عدد مرات الإصابة	التعداد	التكرار	التكرار النسبي	التكرار المئوي
0	||||| |||	8	$8/40 = 0.20$	$0.20\times100 = 20\ \%$
1	||||| ||||| ||||| |||	18	$18/40 = 0.45$	$0.45\times100 = 45\ \%$
2	||||| ||||| ||	12	$12/40 = 0.30$	$0.30\times100 = 30\ \%$
3	||	2	$2/40 = 0.05$	$0.05\times100 = 5\ \%$
Total		**40**	**1**	**100 %**

١-٢-٢-٢- ملاحظات

١- إنَّ مجموع التكرارات النسبيّة يجب أن يساوي الواحد تماماً، ولكن عند تنفيذ بعض الحساب نضطر إلى إجراء عمليّة تدوير للأرقام، وفي هذه الحالة قد لا نحصل على مجموعٍ يساوي الواحد تماماً، فيكون المجموع أكبر أو أصغر من الواحد بقليل.

٢- إنَّ مجموع التّكرارات المئويّة يجب أن يساوي المئة تماماً، ولكن إذا ما حصلت عمليّة تدوير للأرقام فإنَّ مجموع التكرارات المئوية قد لا يساوي المئة تماماً، فيكون لدينا مجموع أكبر أو أصغر من المئة بقليل.

٣- بعد الانتهاء من صبِّ البيانات يمكن الاستغناء عن عمود التعداد لأنَّ عمود التكرار يؤدي الغاية نفسها، وأمَّا في حال تقديم البيانات مجمَّعة في جدول تكراري فإنَّه (وفي معظم الحالات) لا يدرج عمود التعداد معه بسبب عدم وجود مبرر لظهوره، وبذلك يتبقى لدينا جدول بأربعة أعمدة فقط.

نوع آخر من التمثيلات الجدولية للبيانات تقدِّمها لنا الفقرة الآتية.

١-٢-٢-٣- جدول الساق والأوراق Stem and Leafs Table

توجد طريقة أخرى لصب البيانات الخام في جدول مشابه للجداول التكرارية وتُدعى **عرض السّاق والأوراق**، والغاية منـه تصنيف البيانات في مجموعات جزئيــة، كأن نبيّن القيم ما بين 0 و9 في مجموعة جزئية واحدة، والقيم ما بين 10 و19 في مجموعة جزئية ثانية، والقيم ما بين 20 و29 في مجموعة جزئية ثالثة، وهكذا دواليك.

إنَّ صب البيانات الخام في جدول السّاق والأوراق يتمَّ من خلال بناء جدول بعمودين أحدهما يخصَّص للساق Stem والآخر للأوراق Leafs، وذلك على النحو الآتي:

بفرض أنَّه لدينا بيانات كميّة مكوَّنةً من أعداد صحيحة من خانتين على الأكثر، فعندئذٍ نضع في عمود الساق القيمة 0، وفي العمود المقابل لها (عمود الأوراق) نضع جميع القيم المكوَّنة من خانة واحدة فقط. بعد ذلك ننتقل إلى الأرقام المكوَّنة من خانتين فنبدأ بالقيم التي خانة العشرات لها هي الأصغر، فنضع العدد المكوِّن لخانة العشرات في عمود الساق وأمَّا الأجزاء الأخرى المكوِّنة لخانة الآحاد من العدد نفسه فنضعها في عمود الأوراق، فإذا ما انتهينا من هذه ننتقل إلى الأرقام الأكبر والمكوَّنة من خانتين فنقوم بتطبيق الإجراء السابق عليها أيضاً، وهكذا دواليك حتى الانتهاء من عملية الصب. هذا ويفضَّل ترتيب القيم في عمود الأوراق تصاعدياً في العرض الأخير للبيانات، والمثال الآتي يوضِّح لنا ذلك.

١-٢-٢-٤- أمثلة

لتكن لدينا البيانات الآتية التي تمثِّل درجات اختبار المنتصف لـ 24 طالباً:

12	6	21	7	17	14	11	8	22	27	3	13
13	19	5	24	6	15	23	25	25	5	6	16

عندئذٍ باستخدام الطّريقة التي تمَّ شرحها أعلاه أنَّ جدول الساق والأوراق للبيانات المُعطاة كما يلي:

الجدول [1-3-a]

Stem الساق	Leafs الأوراق
0	6 , 7 , 8 , 3 , 5 , 6 , 5 , 6
1	2 , 7 , 1 , 3 , 3 , 9 , 5 , 6
2	1 , 2 , 4 , 7 , 4 , 3 , 5 , 5

الأوراق حسب موضعها في البيانات المقدَّمة

وبعد ترتيب البيانات تصاعدياً يصبح للجدول السابق العرض الآتي:

الجدول [1-3-b]

Stem الساق	Leafs الأوراق
0	3 , 5 , 5 , 6 , 6 , 6 , 7 , 8
1	1 , 2 , 3 , 3, 5 , 6 , 7 , 9
2	1 , 2 , 3 , 4 , 4 , 5 , 5 , 7

في الواقع هذه ليست كلّ التمثيلات الجدولية للبيانات حيث يوجد نوع آخر من التمثيلات الجدولية للبيانات يسمَّى **جدول التوزيع التكراري** وسنأتي على شرحه لاحقاً في هذا الفصل.

١-٢-٣- التّمثيل الشّرائطي للبيانات الخام

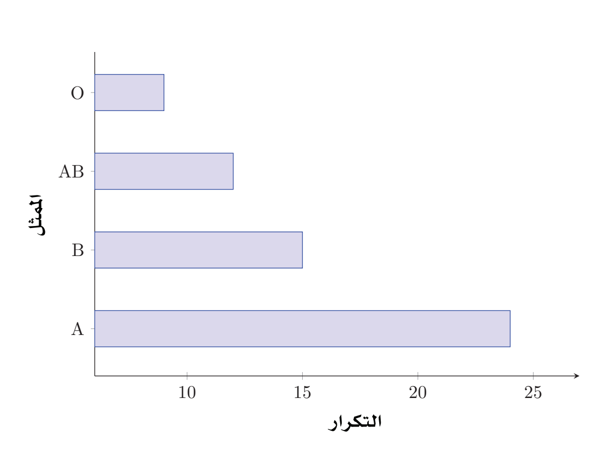
Bar Chart Representation of Raw Data

يعدّ تمثيل البيانات الخام باستخدام الشّرائط العمودية (**ويُعرف باسم التمثيل بالأعمدة أيضاً**) أو الشرائط الأفقية من أحد التمثيلات الجيّدة للبيانات الخام، وذلك لأنَّها تعطي انطباعاً سريعاً حول طبيعة البيانات الخام وسلوكها، وسبب ذلك أنَّه من طبيعة الإنسان سرعة الملاحظة عند النظر إلى المشاهد والرسومات.

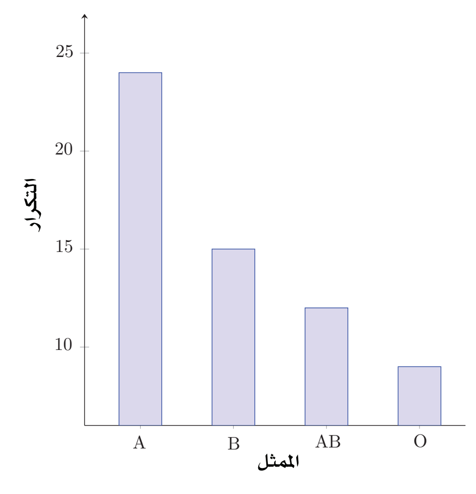
من أجل التمثيل الشّرائطي لمجموعة بيانات خام (**وبغض النظر عن نوعها اسميّةً، رموزاً أو عددية**) نقوم برسم محورين متعامدين **XoY**، ومن ثمَّ يُدوَّن أسفل المحور **oX** المُمثَّل لكل صنف في البيانات الخام (**فإن كانت أسماءً كُتِبَت الأسماء، وإن كانت رموزاً وضِعَت الرموز وإن كانت أعداداً سُجِّلت الأعداد**)، وأمّا المحور **oY** فيدوَّن عليه قيم تكرارات البيانات الخام. بعد ذلك يُرسم عمود فوق كل مُمثِّل بيانات بارتفاع قدره يساوي قيمة تكرار هذا الممثِّل مع الأخذ بالحسبان أن تكون هذه الأعمدة منفصلة بعضها عن البعض الآخر بتباعد ثابت (**وغالباً ما تؤخذ بمقدار وحدة قياس**). في هذه الحالة نحصل على التمثيل بالشرائط العمودية أو التمثيل بالأعمدة. أمّا إذا أردنا تمثيل البيانات بالشرائط الأفقية فإنّنا نعكس العمليات التي تمت على المحورين المتعامدين، فيصبح المحور **oY** من أجل تدوين الممثِّل لكل صنف في البيانات الخام في حين يستخدم المحور **oX** لتدوين قيم تكرارات البيانات الخام، وفي هذه الحالة لا يُقال عن التمثيل الناتج إنَّه تمثيل بالأعمدة للبيانات الخام. بعد ذلك يُرسم شريط أفقي إلى جانب كل مُمثِّل للبيانات بطول قدره يساوي قيمة تكرار هذا الممثِّل، ومع الأخذ بالحسبان أن تكون هذه الأشرطة منفصلة بعضها عن البعض الآخر بتباعد ثابت أيضاً.

١-٢-٣-١- مثال

بالعودة إلى المثال (٢) من (١-٢-٢-١) فإنّنا نجد تمثيل البيانات الخام المُعطاة باستخدام الشّرائط العموديّة (**أو التمثيل بالأعمدة**) له الشكل [1-3-a]، وأمّا التمثيل باستخدام الشّرائط الأفقيّة فله الشكل [1-3-b].



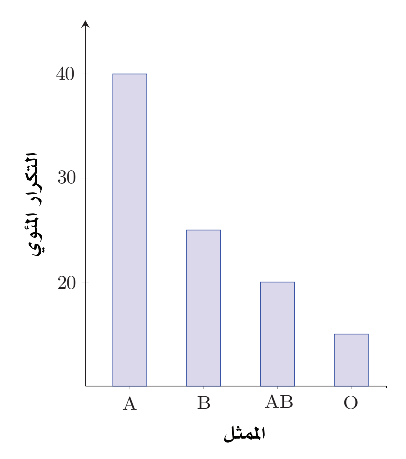
الشكل [1-3-b]



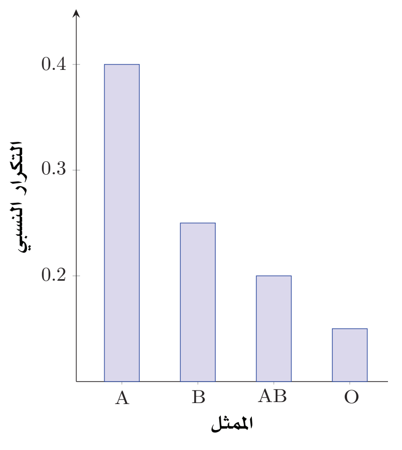
الشكل [1-3-a]

١-٢-٣-٢- ملاحظة

يمكن استخدام التكرارات النّسبيّة والتكرارات المئوية بدلاً من التكرارات في التمثيل الشرائطي أيضاً، حيث تستبدل قيم التكرارات بقيم التكرارات النسبيّة أو التكرارات المئوية، فعلى سبيل المثال نجد من أجل المثال السابق أنَّ العرض الشرائطي النسبي والمئوي لهما الشكلين الآتيين:



الشكل [1-3-d] العرض الشرائطي المئوي



الشكل [1-3-c] العرض الشرائطي النسبي

١-٢-٣-٣- التّمثيل بالشّرائط البيانيّة المزدوجة Pair Bar Charts Representation

يُعدّ التمثيل البياني بالشرائط المزدوجة من التمثيلات البيانية المهمّة عند مقارنة مجموعتين من البيانات أو أكثر (ظاهرتين أو أكثر) مع بعضها البعض الآخر، وهذه الطريقة في التمثيل لها الخطوات نفسها التي استخدمت من أجل التمثيل الشرائطي، ولكن لكل مجموعة بيانات على حدى مع وضع الأشرطة الممثِّلة للنوع الواحد ملاصقة أو قريبة بعضها من البعض الآخر، ويُميَّز أحدهما عن الآخر بالتّلوين المختلف أو التظليل المختلف.

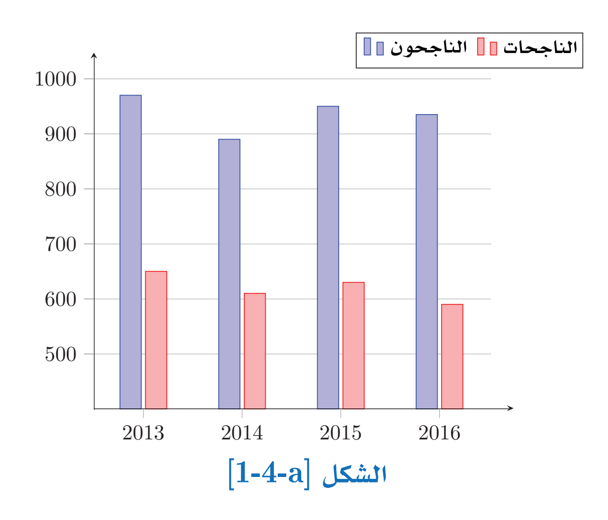
١-٢-٣-٤- مثال

على سبيل المثال يمكننا أن نقارن أعداد الطلاب والطالبات الذين اجتازوا اختبار مقرَّر الإحصاء في السنة الأولى المشتركة في جامعة الملك سعود خلال الأعوام الدراسية 2012-2013 وحتى 2015-2016 حيث لدينا البيانات الخاصّة بذلك كما في الجدول الآتي:

الجدول [1-4]

العام الدراسي	عدد الطلاب الناجحون	عدد الطالبات الناجحات
2012-2013	970	650
2013-2014	890	610
2014-2015	950	630
2015-2016	935	590
Total	**3745**	**2480**

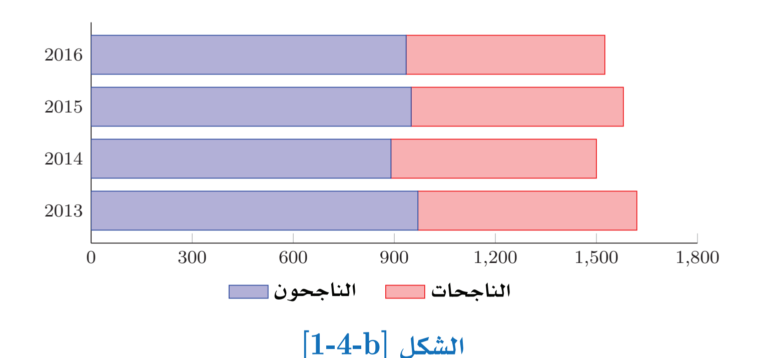
فعندئذٍ نجد أنَّ التمثيل بالشرائط المزدوجة للبيانات المقدَّمة أعلاه له الشكل الآتي.



الشكل [1-4-a]

١-٢-٣-٥- التّمثيل بالشّرائط المجزّأة Stacked Bar Charts Representation

يعدّ التمثيل البياني باستخدام الشرائط المجزأة إحدى الطرائق المستخدمة في مقارنة مجموعتين من البيانات أو أكثر (ظاهرتين أو أكثر)، ويتمّ ذلك بطريقة مماثلة لما سبق في طريقة العرض الشرائطي المزدوج، ولكن بدلاً من أن توضع الشرائط بعضها إلى جانب البعض الآخر، فإنَّها توضع بعضها فوق البعض الآخر، ولهذا السّبب يجب مراعاة وضع قيم التكرارات بحيث تغطي مجموع التكرارات لكل ممثِّل من البيانات. هذا ويفضل استخدام هذه الطريقة في التمثيل عندما نكون مهتمين بالتكرارات الكليّة لممثِّلي البيانات في مجموعات البيانات المُعطاة، فعلى سبيل المثال نجد التمثيل بالشرائط المجزأة لبيانات المثال (١-٢-٣-٤) له الشكل الآتي:



الشكل [1-4-b]

١-٢-٤- التّمثيل بالقطاعات الدّائريّة (أو القرص الدائري) Pie Chart

يُنظر إلى التمثيل بالقطاعات الدّائريّة (أو القرص الدّائري) على أنَّه أحد الأشكال البيانية الواسعة الانتشار في دراسات الإحصاء الوصفي عندما يكون عدد ممثّلي البيانات قليلاً، وأمَّا إذا كان عدد ممثّلي البيانات كبيراً فعندئذٍ تصبح الفائدة منه شبه معدومة.

يُلجأ عادة إلى استعمال هذه الطريقة عندما نكون بحاجة لتقسيم الكل إلى k من الأجزاء، وأمَّا لرسمها فإنَّنا نقوم أولاً برسم دائرة بنصف قطر مُثبَّت (غالباً يكون عمودياً) يُعدّ مبدأً لقياس الزّاوية عنه، ثمَّ تُحسب زوايا القطاعات الدائرية α_i مقدَّرة بالدّرجات Degrees وتأخذ إلى يمين العمود السابق باتجاه دوران عقارب السّاعة، وبحيث يكون للممثِّل (أو الجزء) i قطّاع دائري زاويته تُحسب بوساطة العلاقة الآتية:

$$\alpha_i := \frac{n_i}{n} \times 360 \qquad [1-1]$$

علماً أنَّ n هو عدد البيانات الخام المُعطاة و n_i هو عدد العناصر (أو البيانات) التابعة للممثِّل (أو الجزء) i، بمعنى آخر، فإنَّنا نحصل على قيمة الزّاوية للقطاع التابع للممثِّل (أو الجزء) i من خلال ضرب قيمة التكرار النسبي لهذا الممثِّل في 360، والمثالان الآتيان يوضّحان لنا ذلك.

١-٢-٤-١ مثال

١- بالرّجوع إلى المثال (٢) من (١-٢-٢-١) وباستخدام العلاقة [1-1] نجد أنَّ:

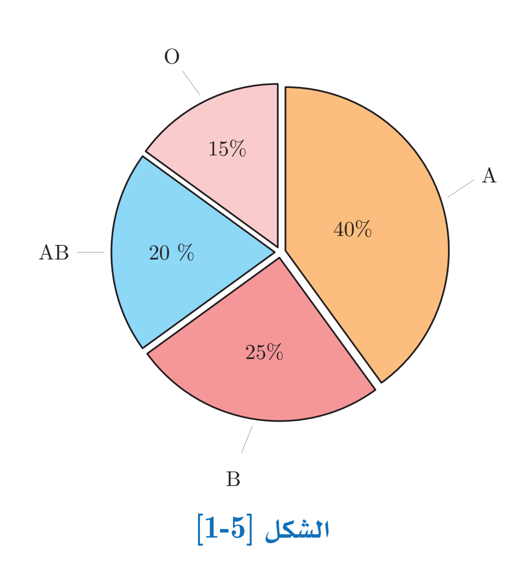
زاوية القطاع الدائري لممثِّل البيانات A هي $\alpha_A = \frac{24}{60} \times 360 = 144^o$

زاوية القطاع الدائري لممثِّل البيانات B هي $\alpha_B = \frac{15}{60} \times 360 = 90^o$

زاوية القطاع الدائري لممثِّل البيانات AB هي $\alpha_{AB} = \frac{12}{60} \times 360 = 72^o$

زاوية القطاع الدائري لممثِّل البيانات O هي $\alpha_O = \frac{9}{60} \times 360 = 54^o$

ومن ثمَّ يكون لدينا العرض الآتي للقطاعات الدائرية.



الشكل [1-5]

٢- بالرجوع إلى المثال (٣) من (١-٢-٢-١) وباستخدام العلاقة [1-1] نجد أنَّ:

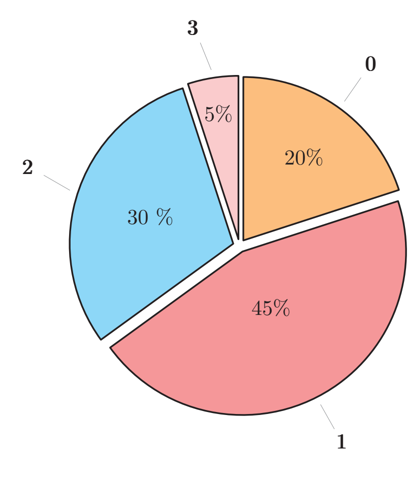
زاوية القطاع الدائري لممثِّل البيانات 0 هي $\alpha_0 = \frac{8}{40} \times 360 = 72^o$

زاوية القطاع الدائري لممثِّل البيانات 1 هي $\alpha_1 = \frac{18}{40} \times 360 = 162^o$

زاوية القطاع الدائري لممثِّل البيانات 2 هي $\alpha_2 = \frac{12}{40} \times 360 = 108^o$

زاوية القطاع الدائري لممثِّل البيانات 3 هي $\alpha_3 = \frac{2}{40} \times 360 = 18^o$

ومن ثمَّ يكون لدينا العرض الآتي للقطاعات الدائرية.



الشكل [1-6]

١ - ٣
التوزيعات التكرارية

لقد قدّمنا فيما سبق شرحاً مفصّلاً للجداول التكرارية حيث لاحظنا أنَّ تلك الجداول تعطينا تصوّراً سريعاً حول سلوك البيانات، وأكثر من ذلك فقد كانت تهتمّ بسلوك ممثِّل البيانات نفسه، فتظهر لنا تكراره وتكراره النسبي والمئوي إذا رغبنا في ذلك. إلاَّ أنَّه من أجل البيانات الكميّة المستمرّة (أو المتّصلة) خصوصاً فقد لا نكون قادرين على استخدام هذه الطريقة في العرض بسبب أنَّ قيمها تنتمي إلى مجموعات قد تكون غير قابلة للعدّ، وربما لا تكرَّر قيمة البيان لأكثر من مرَّةٍ واحدة أيضاً. لذلك في مثل هذه الحالات لا يعود اهتمامنا مركّزاً على قيمة البيان نفسه وإنّما على الفترة التي ينتمي إليها هذا البيان، ومن ثمَّ يصبح اهتمامنا منصباً على بيان ممثِّل لكلِّ فترةٍ من الفترات التي ستصبّ فيها البيانات. بمعنى آخر، يكون لدينا تجزئة لمجموعة البيانات في فترات جزئية واهتمامنا يكون منصباً على قيمة ممثِّلة وحيدة (**تُدعى مركز الفئة – سيرد ذكرها لاحقاً-**) لكل فترة من هذه الفترات الجزئية. عادة يُطلق اسم "**فئة** Class" على كل فترة جزئية من هذه الفترات.

إنَّ تمديد الجداول التكرارية إلى جداول تكرارية ذوات فئات تصبح أكثر فاعلية في تنفيذ بعض العمليات الحسابية على البيانات وخاصّة عندما يصبح عدد البيانات كبيراً، فعلى سبيل المثال تصوّر لو أنَّك تقوم بدراسة على مجموعة بيانات مكوَّنة من مليون قيمة عددية أو أكثر فكم من الوقت ستحتاج لاستنباط سلوك هذه البيانات أو الحصول على بعض المميّزات العددية لها؟

قبل البدء في بناء جداول التوزيع التكرارية لا بدَّ لنا من تقديم بعض المفاهيم التي لا بدَّ منها لبناء مثل هذه الجداول، وسنبدأها بالتعريف الآتي.

١-٣-١- تعريف (المدى Range)

لتكن لدينا x_1، x_2، ... و x_n بيانات مُعطاة، ولنرمز لأصغر وأكبر قيمة في هذه البيانات بـ x_s و x_ℓ على الترتيب. عندئذٍ يُعرَّف **المدى** لهذه البيانات (**وسنرمز له بـ R**) على أنَّه الفرق بين أكبر وأصغر قيمة في هذه البيانات. أي أنَّه لدينا:

$$\mathbf{R} = x_\ell - x_s \qquad [1\text{-}2]$$

١-٣-٢- تعريف (الفئة Class)

الفئة من أجل بيانات مُعطاة هي فترة من مجموعة الأعداد الحقيقية لها طول موجب تماماً وتحتوي على بعضٍ من البيانات المُعطاة، ويقال عن طرفها الأيسر إنَّه الحد الأدنى للفئة في حين يُقال عن طرفها الأيمن إنَّه الحد الأعلى للفئة.

من التّسميات الأخرى للفئة Interval أو Category أو Group.

١-٣-٣- بناء جدول التوزيع التكراري

بالرجوع إلى موضوع تمديد الجداول التكرارية فإنَّ عملية تمديد هذه الجداول وفقاً للآلية التي سنقدِّمها بعد قليل تُعطينا ما يُعرف باسم "جداول التوزيع التكرارية"، ولهذه الجداول نماذج مختلفة ولكن معظم هذه الجداول تحتوي على الأعمدة المقدَّمة في الجدول الآتي:

الجدول [1-5]

التكرار المتجمِّع الصاعد للفئة	التكرار المئوي للفئة	التكرار النسبي للفئة	تكرار الفئة	تعداد الفئة	مركز الفئة	الحدود الفعليّة للفئة	الحدود العمليّة للفئة	رقم الفئة
...	...	...	...	...	...	...	...	1
...	...	...	...	...	...	...	...	2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
المجموع				-------	-------	--------	--------	Total

وبخصوص نوع الفئات لجداول التوزيع التكرارية يمكن أن تُعرض (أو تُقدَّم) وفقاً لأحد نوعين من الفئات:

أ- جداول توزيع تكرارية تحتوي على فئات ذات أطوال مختلفة، وهذا النوع لن نقوم بدراسته في هذا الكتاب.

ب- جداول توزيع تكرارية تحتوي على فئات ذات أطوال متساوية، وهذا النوع من الجداول سيكون محور دراستنا في هذا الكتاب.

قبل البدء في كيفية بناء جدول التوزيع التكراري نودّ التنويه إلى أنَّنا سنشرح بناء جدول التوزيع التكراري من أجل الحالات البسيطة التي تكون فيها قيمة المدى للبيانات كبيرة نسبياً، والأمثلة التي سنقدّمها ستكون على قيم صحيحة للبيانات، وكذلك لن نتطرق إلى الحالات التي تستوجب معالجة خاصة في بناء جدول التوزيع التكراري.

الآن، ومن أجل بناء جدول توزيع تكراري نتبع الخطوات الآتية:

١- تعيين عدد فئات جدول التوزيع التكراري:

إذا قدِّم لنا عدد الفئات (وليكن k فئة) الواجب استخدامها في **جداول التوزيع التكراري** من قبل الجهة الطالبة لدراسة المسألة الإحصائية، فإنّنا نلتزم بهذا العدد للفئات ولا نجري أيّ تعديل عليه إلاَّ لضرورة تتطلّبها شروط بناء جداول التوزيع التكرارية. لكن إذا لم يقدَّم عدد الفئات الواجب دراستها فإنَّنا نختار عدد الفئات k بحيث لا يقلُّ عن خمس فئات ولا يزيد عن عشرين فئة، وذلك لأنَّ جداول التوزيع التكرارية التي تحتوي على أقلّ من خمس فئات تعدّ قليلة الفائدة.

وأمّا التي تحتوي على أكثر من عشرين فئة فإنّها تكون متعبة في الدراسة، وهذه التّوصية تتحقّق عمليّاً باستخدام العلاقة الآتية (مستخدمة في بعض البرامج الإحصائية) لحساب عدد الفئات k ما دام عدد المشاهدات (البيانات) n أكبر أو يساوي 32 مشاهدة (بطبيعة الحال من غير المرغوب تجميع البيانات في جداول توزيع تكرارية إذا كان عددها أقل من 30):

$$k = \lfloor 3.322 \log n \rfloor \qquad [1-3]$$

علماً أنَّ $\lfloor x \rfloor$ هو أكبر عدد صحيح أصغر أو يساوي x ، فعلى سبيل المثال $\lfloor 5 \rfloor = 5$ وكذلك لدينا:

$$\lfloor 3.322 \log 32 \rfloor = \lfloor 5.00011 \rfloor = 5 \quad \& \quad \lfloor 3.322 \log 31 \rfloor = \lfloor 4.9543 \rfloor = 4$$

$$\lfloor 3.322 \log 64 \rfloor = \lfloor 6.00013 \rfloor = 6 \quad \& \quad \lfloor 3.322 \log 63 \rfloor = \lfloor 5.97741 \rfloor = 5$$

٢- تعيين سعة وحدود الفئات في جدول التوزيع التكراري:

بفرض أنَّه لدينا بيانات عددها n ولها مدى $\mathbf{R}$ وبعدد فئات k، فعندئذٍ نحصل على سعة الفئة الفعليّة Length of Class Boundary (ونرمز لها بـ $\mathbf{C}$) باستخدام العلاقة الآتية:

$$\mathbf{C} = \frac{\mathbf{R}}{k} \qquad [1-4]$$

وننوّه هنا إلى أنَّ سعة الفئة النّاتجة عن الحساب السّابق يمكن تقريبها بالزّيادة قليلاً إلى قيمة أكبر بحيث تسمح لنا القيمة الجديدة للسعة بتفريغ أسهل للبيانات، ولكن يجب عدم المبالغة في الزيادة لأنَّ الزيادة المبالغ فيها قد تؤدي إلى توليد فئات في آخر جدول التوزيع التكراري بحيث يكون حدُّها الأدنى خارج نطاق البيانات المُعطاة، ومن ثمَّ سيكون تكرارها معدوماً (ستصبح فئة خالية من البيانات).

- تعيين حدود الفئات الفعليّة (أو الحقيقية) Class Boundary:

الآن، وبعد تحديد السعة لكل الفئات الفعليّة $\mathbf{C}$، فإنّنا نقوم بتعيين حدود هذه الفئات كما يلي:

أ - نجعل الحد الأدنى لأول فئة فعليّة مساوياً للقيمة الناتجة عن طرح 0.5 من أصغر قيمة في البيانات.

ب- نضيف قيمة $\mathbf{C}$ إلى الحد الأدنى لهذه الفئة فنحصل على الحد الأعلى للفئة الفعليّة الأولى.

ج- نجعل الحد الأدنى للفئة الفعليّة التالية (الثانية) مساوياً للحد الأعلى للفئة الفعليّة السابقة (الأولى)، ومن ثمَّ نضيف قيمة $\mathbf{C}$ إلى الحد الأدنى لهذه الفئة فنحصل على الحد الأعلى للفئة الفعليّة الثانية.

د- نقوم بتطبيق الفقرة السابقة (ج) من أجل جميع الفئات المتبقية فنحصل على الحدود الدنيا والعليا للفئات الفعليّة لجدول التوزيع التكراري.

- تعيين حدود الفئات العمليّة (أو التّجريبيّة) Class Limit:

يُعيَّن الحد الأدنى للفئة العمليّة ذات الرقم i مع $i = 1,2,...,k$ من خلال إضافة 0.5 إلى الحد الأدنى للفئة الفعليّة ذات الرقم i (لاحظ هنا أنَّ الحد الأدنى لأول فئة عمليّة سيوافق أصغر قيمة في البيانات المُعطاة)، وأمَّا الحد الأعلى للفئة العمليّة ذات الرقم i فإنَّه يُعيَّن من خلال طرح 0.5 من الحد الأعلى للفئة الفعليّة ذات الرقم i.

تجدر الإشارة هنا إلى الملاحظات الآتية:

١- أنَّ عملية تفريغ البيانات في جدول توزيع تكراري تتمّ في الفئات الفعليّة حصراً لأنَّه لدى عملية تحصيل البيانات سيكون لدينا خطأ مرتكب من وسائل القياس المعتمدة (مهما بلغت من دقّة) قد لا تصل إلى القيمة الحقيقية لقياس المشاهدة، ولذلك تمّ الاتفاق على أنَّ نصف وحدة الدقّة المعتمدة في القياس ستغطي هذا الخطأ زيادةً أو نقصاناً، بمعنى أنَّه بهذه العملية سيُغطي أكبر خطأ محتمل لدى أخذ البيانات، ولهذا السبب سمّيت هذه الفئات بالفئات الفعليّة.

٢- إذا وافقت قيمة x من قيم البيانات الحد الأدنى لفئة فعليّة فإنَّها تفرغ في هذه الفئة، وأمّا إذا وافقت هذه القيمة الحد الأعلى للفئة الفعليّة فإنَّها تفرغ في الفئة الفعليّة التالية، ولهذا السبب سنكتب (وعلى سبيل التوضيح) الفئة الفعليّة التي حدّها الأدنى a وحدّها الأعلى b بالشكل $a \rightarrow b$ للدلالة على أنَّ القيمة b لا تتبع هذه الفئة وإنَّما تتبع الفئة الفعليّة التالية.

٣- وفقاً لاستخدام العلاقة [4-1] في تعيين سعة الفئة قد يحصل أنَّ قيمة x أو أكثر من قيم البيانات تكون أكبر أو تساوي الحد الأعلى للفئة الفعليّة الأخيرة، فإذا وافقت القيم المتبقية الحد الأعلى للفئة الفعليّة الأخيرة فإنَّنا نقوم بتحميلها في الفئة الفعليّة الأخيرة، وأمّا إذا كان هناك قيم أكبر من الحد الأعلى للفئة الفعليّة الأخيرة فإنَّنا نقوم بزيادة السعة قليلاً بحيث تحتوي الفئة الفعليّة الأخيرة باقي البيانات.

٤- من أجل فئة فعليّة حدّها الأدنى a وحدّها الأعلى b ستكون للفئة العمليّة الموافقة حدّ أدنى α قيمته تزيد على a بمقدار نصف وحدة الدّقة المعتمدة، وأمّا حدّها أعلى β فإنَّه سينقص عن b بمقدار نصف وحدة الدّقة المعتمدة، ولذلك تُكتب الفئة العمليّة بالشكل $\alpha - \beta$ (بدون سهم $\rightarrow$) للدلالة على أنَّه من أجل هذا النوع من الفئات يكون كل من الحدّ الأدنى والأعلى منتمياً لها.

٥- عندما يقدَّم جدول توزيع تكراري لمجموعة بيانات (أي أنَّ البيانات الخام التي صُبَّت فيه ليست ظاهرة - خفية-) فإنَّه يمكن الاستغناء عن العمود الخاص بالفئات العمليّة لعدم الحاجة له في أية دراسة لاحقة تتعلَّق بهذا الجدول.

٦- إذا حصل لدينا فئات فعليّة خالية من البيانات عند عمليّة التفريغ فإنَّنا نقوم بتغيير عدد الفئات (وغالباً بزيادة عددها) حتى تختفي جميع الفئات الخالية. بالطبع سيرافق هذا التغيير لعدد الفئات تغيُّر في سعتها أيضاً، ولذلك يجب الانتباه في حال التعديل على السعة الجديدة ألاَّ يؤدي ذلك إلى فئات خالية من البيانات أيضاً.

٤- تعيين مراكز الفئات Class Centers:

إنَّ مركز الفئة يساوي نصف مجموع حدّيها الأعلى والأدنى (أية فئة كانت العمليّة أو الفعليّة)، ويُنظر إلى هذه القيمة (أي إلى مركز الفئة) على أنَّها الممثِّل لكل البيانات التي ستنتمي إلى هذه الفئة، ولذلك سوف نلاحظ أنَّ لهذه القيمة دوراً مهمّاً جداً لدى حساب القيم العددية المميِّزة للبيانات المجمّعة في جداول توزيع تكرارية.

٥- تدوين التّعدادات للفئات Class Tallies:

لقد قمنا سابقاً بشرح تدوين التعدادات من أجل الجداول التكرارية، وهنا يستخدم بآليّة مماثلة، ولكن هنا يرسم خط من أجل كل قيمة بيان تنتمي إلى الفئة الفعليّة (وليس إلى الفئة العمليّة).

نشير هنا إلى أنَّ هذا العمود يستخدم عندما يكون لدينا عمليّة تفريغ لبيانات خام (بيانات كميّة) في جدول توزيع تكراري، فيما عدا ذلك لا ضرورة لهذا العمود ضمن جدول التوزيع التكراري ويمكن حذفه.

٦- تعيين قيم التكرارات للفئات Frequencies of Classes:

لقد قمنا سابقاً بشرح تعيين قيم التكرار للفئات من أجل الجداول التكرارية، ويستخدم من أجل جداول التوزيع التكرارية بالآلية نفسها.

٧- تعيين قيم التكرارات النسبيّة للفئات Relative Frequencies of Classes:

لقد قمنا سابقاً بشرح تعيين قيم التكرار النسبيّة للفئات من أجل الجداول التكرارية، ويستخدم من أجل جداول التوزيع التكرارية وفقاً للآلية نفسها أيضاً.

٨- تعيين قيم التكرارات المئوية للفئات Percentile Frequencies of Classes:

لقد قمنا سابقاً بشرح تعيين قيم التكرار المئوية للفئات من أجل الجداول التكرارية، ويستخدم من أجل جداول التوزيع التكرارية وفقاً للآليّة نفسها أيضاً.

٩- تعيين قيم التكرارات المتجمِّعة الصاعدة للفئات:
Ascending Cumulative Frequencies of Classes

العمود التالي والأخير من جدول التوزيع التكراري هو عمود التكرارات المتجمِّعة الصاعدة، وتدوَّن القيم في هذا العمود على النّحو الآتي:

مقابل الفئة الأولى يدوَّن تكرار الفئة الأولى نفسه (ويُدعى التكرار المتجمِّع الصاعد للفئة الأولى) لأنَّه يمثِّل كل التكرارات التي أقل من حدِّها الأعلى. أمَّا مقابل الفئة الثانية فيتمّ تدوين حاصل مجموع التكرارين للفئتين الأولى والثانية (ويُدعى التكرار المتجمِّع الصاعد للفئة الثانية) وهو يمثِّل كلّ التكرارات التي أقلّ من حدِّها الأعلى. يدوَّن مقابل الفئة الثّالثة حاصل مجموع التكرارات للفئات الأولى والثانية والثالثة (ويُدعى التكرار المتجمِّع الصاعد للفئة الثالثة) وهو يمثِّل كلّ التكرارات التي أقلّ من حدِّها الأعلى، وهكذا دواليك حتى الانتهاء من جميع الفئات. إنَّ القيم المدوَّنة في هذا العمود تُدعى التكرارات المتجمِّعة الصّاعدة.

١-٣-٤- مثال

البيانات الآتية تمثِّل عدد الأميال المقطوعة لكلّ لتر من البنزين لأربعين سيارة حديثة.

13	14	16	20	17	19	12	15	16	12
18	12	15	16	14	20	15	12	20	12
12	14	16	18	20	12	14	16	18	20
14	15	16	15	18	16	15	16	17	16

ولنقم بتفريغ هذه البيانات في جدول توزيع تكراري، ولكن بسبب كِبر نموذج جدول التوزيع التكراري (كجدول نموذجي للتدريب) سنقوم (وعلى سبيل التوضيح) بتجزئة الجدول إلى جدولين.

من أجل ذلك لدينا أكبر قيمة في البيانات $x_\ell = 20$، وأصغر قيمة في البيانات $x_s = 12$، ومن ثمّ نجد باستخدام العلاقة [1-2] أنَّ المدى للبيانات المُعطاة يساوي:

$$\mathbf{R} = x_\ell - x_s = 20 - 12 = 8$$

والآن لتعيين عدد الفئات نلاحظ أنَّه لدينا 40 مشاهدة، ومن ثمَّ باستخدام العلاقة [1-3] يكون عدد الفئات اللازمة لجدول التوزيع التكراري المطلوب هو:

$$k = \lfloor 3.\ 322 \log 40 \rfloor = \lfloor 5.322 \rfloor = 5$$

ومن ثمَّ بحسب العلاقة [1-4] تكون سعة الفئة لجدول التوزيع التكراري المطلوب هي:

$$\mathbf{C} = \frac{\mathbf{R}}{k} = \frac{8}{5} = 1.6$$

من أجل اختيار السّعة المناسبة للفئات الفعليّة لنمعِّن النظر في الجدول الآتي:

الجدول [1-6-a]

من أجل C=2		من أجل C=1.6	
الحد الأعلى للفئة	الحد الأدنى للفئة	الحد الأعلى للفئة	الحد الأدنى للفئة
13.5	11.5	13.1	11.5
15.5	13.5	14.7	13.1
17.5	15.5	16.3	14.7
19.5	17.5	17.9	16.3
21.5	19.5	19.5	17.9

لاحظ أنَّه لو استخدمنا السعة $\mathbf{C} = 1.6$ للفئات فإنَّ الحد الأعلى للفئة الفعليّة الأخيرة سيكون مساوياً لـ 19.5 ومن ثمَّ القيمة 20 في البيانات المُعطاة ستبقى بدون تفريغ، ولذلك سنقوم بزيادة سعة الفئة إلى القيمة 2 (حيث تصبح عملية تفريغ البيانات أبسط) فنجد أنَّ الحد الأعلى للفئة الفعليّة الأخيرة سيكون مساوياً لـ 21.5، وبالتالي نضمن تفريغ جميع البيانات في فئات جدول التوزيع التكراري.

الآن باتّباع الخطوات التي قمنا بشرحها سابقاً فإنَّنا سنميِّز التفريغ للبيانات باستخدام لون خاص لكل فئة (وذلك على سبيل التّوضيح فقط) فيكون لدينا العرض الآتي للبيانات المُعدّة للتّفريغ:

13	14	16	20	17	19	12	15	16	12
18	12	15	16	14	20	15	12	20	12
12	14	16	18	20	12	14	16	18	20
14	15	16	15	18	16	15	16	17	16

وبتفريغ هذه البيانات سنحصل على جدول التوزيع التكراري الآتي الذي يحتوي على الفئات الفعليّة، الفئات العمليّة، مراكز الفئات، التعدادات والتكرارات:

الجدول [1-6-b]

تكرار الفئة	تعداد الفئة	مركز الفئة	الحدود الفعليّة للفئة	الحدود العمليّة للفئة	رقم الفئة
8	卌 |||	12.5	$11.5 \rightarrow 13.5$	$12 - 13$	1
10	卌 卌	14.5	$13.5 \rightarrow 15.5$	$14 - 15$	2
12	卌 卌 ||	16.5	$15.5 \rightarrow 17.5$	$16 - 17$	3
5	卌	18.5	$17.5 \rightarrow 19.5$	$18 - 19$	4
5	卌	20.5	$19.5 \rightarrow 21.5$	$20 - 21$	5
40	------	-----	------------	--------------	Total

وأمَّا من أجل جدول التوزيع التكراري الذي يحوي التكرار النسبي، المئوي والمتجمّع الصاعد فإنَّنا سنكون بحاجة إلى عمود الفئات الفعليّة والتكرارات، فيكون لدينا الجدول الآتي.

الجدول [1-6-c]

التكرار المتجمّع الصاعد الفئة	التكرار المئوي للفئة	التكرار النسبي للفئة	تكرار الفئة	الحدود الفعليّة للفئة	رقم الفئة
8	$0.20 \times 100 = 20\%$	$8/40 = 0.20$	8	$11.5 \rightarrow 13.5$	1
$8+10=18$	$0.25 \times 100 = 25\%$	$10/40 = 0.25$	10	$13.5 \rightarrow 15.5$	2
$8+10+12=30$	$0.30 \times 100 = 30\%$	$12/40 = 0.30$	12	$15.5 \rightarrow 17.5$	3
$8+10+12+5=35$	$0.125 \times 100 = 12.5\%$	$5/40 = 0.125$	5	$17.5 \rightarrow 19.5$	4
$8+10+12+5+5=40$	$0.125 \times 100 = 12.5\%$	$5/40 = 0.125$	5	$19.5 \rightarrow 21.5$	5
المجموع	100 %	1	40	-------------	Total

١-٣-٥- ملاحظة

بما أنَّ التّعدادات تستخدم عند تفريغ البيانات فقط، فإنَّه إذا كانت قيم التكرارات معلومة فعندئذٍ يمكن استنتاج قيم التكرارات النسبيّة والتكرارات المئوية عند اللزوم، ولذلك تحذف الأعمدة الخاصّة بالتعدادات والتكرارات النّسبيّة والتكرارات المئويّة من جداول التوزيع التكرارية.

كذلك سنلاحظ لاحقاً أنَّ العروض البيانية وحساب القيم العددية المميِّزة لبيانات جداول التوزيع التكرارية تعتمد على الفئات الفعليّة فقط، ولذلك يحذف العمود الخاص بالفئات العمليّة من جداول التوزيع التكرارية مالم تكن هناك عملية تفريغ لبيانات خام في الجدول، فعلى سبيل التوضيح يُعرض جدول التوزيع التكراري لبيانات المثال السابق **بشكله المختزل** على النحو الآتي:

الجدول [1-6-d]

رقم الفئة	الحدود الفعليّة للفئة	مركز الفئة	تكرار الفئة	التكرار المتجمّع الصاعد للفئة
1	11.5 → 13.5	12.5	8	8
2	13.5 → 15.5	14.5	10	18
3	15.5 → 17.5	16.5	12	30
4	17.5 → 19.5	18.5	5	35
5	19.5 → 21.5	20.5	5	40
Total	----------------	--------	40	المجموع

١ - ٤
التمثيلات البيانية لجداول التوزيعات التكرارية

في كثير من الحالات يكون التعامل مع جداول التوزيع التكرارية أمراً شاقاً، وخاصة إذا كان عدد الفئات كبيراً، لذلك يحاول المرء عرض نتائجه من خلال أشكال بيانية مناسبة يَسْهل معها فهم طبيعة وسلوك بيانات هذه الجداول. من النماذج المستخدمة في هذه العروض البيانية النموذج الآتي.

١-٤-١- المدرَّج التكراري Histogram

ينظر إلى المدرجات التكرارية على أنَّها من الأشكال البيانية المفيدة في تمثيل بيانات الجداول التكرارية، ويتميَّز بسهولته وبساطته. يتكوَّن المدرَّج التكراري من مستطيلات متلاصقة ترسم فوق الفئات الفعليّة للبيانات وبحيث يكون ارتفاع كل مستطيل متناسباً مع قيمة التكرار للفئة التي رُسم عليها، حيث تُمثَّل الفئات الفعليّة على المحور الأفقي (المحور oX)، وأمّا على المحور العمودي عليه (أي المحور الرأسي oY) فيتمّ تدوين قيم التكرارات للفئات الفعليّة.

في الواقع يُمكِّننا الرسم البياني للمدرَّج التكراري من معرفة شكل توزيع البيانات وانتشارها وأين تتمركز البيانات بشكل سريع وسهل، والمثال الآتي يوضِّح لنا ذلك.

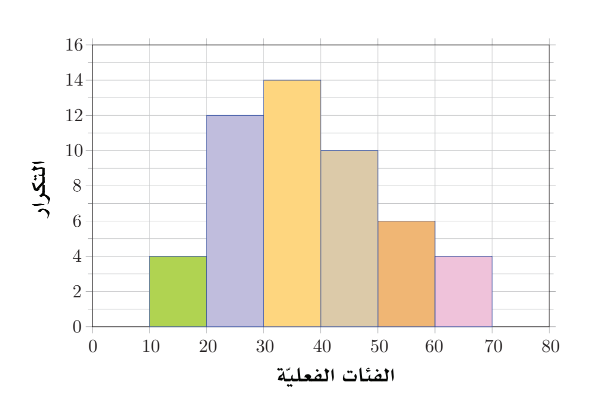
١-٤-١-١- مثال

ليكن لدينا جدول التوزيع التكراري الآتي:

الجدول [1-7]

رقم الفئة	الحدود الفعليّة للفئة	مركز الفئة	تكرار الفئة	التكرار المتجمِّع الصاعد للفئة
1	$10 \to 20$	15	4	4
2	$20 \to 30$	25	12	16
3	$30 \to 40$	35	14	30
4	$40 \to 50$	45	10	40
5	$50 \to 60$	55	6	46
6	$60 \to 70$	65	4	50
Total	-----------	-----	50	المجموع

والآن لرسم المدرَّج التكراري نقوم برسم أعمدة مستطيلة فوق الفئات الفعليّة بحيث يكون ارتفاع كل مستطيل يساوي تكرار الفئة التي رُسم عليها، فنحصل على الشكل [1-7-a] للمدرَّج التكراري.

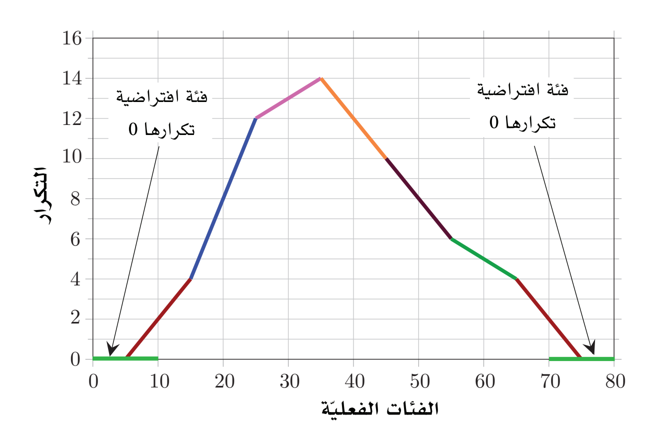


الشكل [1-7-a]

١-٤-٢- المضلَع التّكراري Frequency Polygon

تعدّ المضلّعات التكرارية من الأشكال البيانية المهمّة في تمثيل البيانات الكميَّة أيضاً، وهو رسمٌ بياني مكوَّن من محورين متعامدين حيث تُمثَّل الفئات الفعليّة على المحور الأفقي (المحور oX) في حين تدوَّن قيم التكرارات للفئات الفعليّة على المحور الرّأسي (المحور oY)، ويتمّ تشكيل هذا المضلَّع من خــلال الوصل بقطع مستقيمة بين تلك النقاط التي إحداثياتها على محور الفئات هي مراكز الفئات، وأمّا إحداثياتها على محور التكرارات فهي قيم التكرارات المقابلة لتلك الفئات، وبعد ذلك إغلاق هذا المضلَّع الناتج إلى محور الفئات من خلال وصل بداية المضلَّع الناتج إلى مركز فئة افتراضية (وهمية) سابقة لأول فئة تكرارها معدوم، وبعد ذلك وصل نهاية هذا المضلَّع إلى مركز فئة افتراضية (وهمية) لاحقة بآخر فئة تكرارها معدوم أيضاً.

فعلى سبيل المثال لو رجعنا إلى المثال السابق (١-٤-١-١) نجد أنَّ المضلَّع التكراري لبيانات جدول التوزيع التكراري المُعطى له الشكل الآتي.



الشكل [1-7-b]

أخيراً نشير إلى أنَّه توجد حالات يغلق فيها المضلَّع النّاتج إلى بداية الفئة الأولى ونهاية الفئة الأخيرة إلاَّ أنَّنا لن نتطرق إلى النّقاش الخاص بهذه الحالة.

١-٤-٣- مضلّع التّكرار المتجمِّع الصّاعد Ogive

لرسم مضلَّع التّكرار المتجمِّع الصّاعد نقوم برسم محورين متعامدين، ومن ثمَّ يُمثَّل على المحور الأفقي الفئات الفعليّة في حين يُمثَّل على المحور الرّأسي التكرارات المتجمِّعة الصّاعدة، وبعد ذلك تُعيَّن مجموعة نقاط على النحو الآتي:

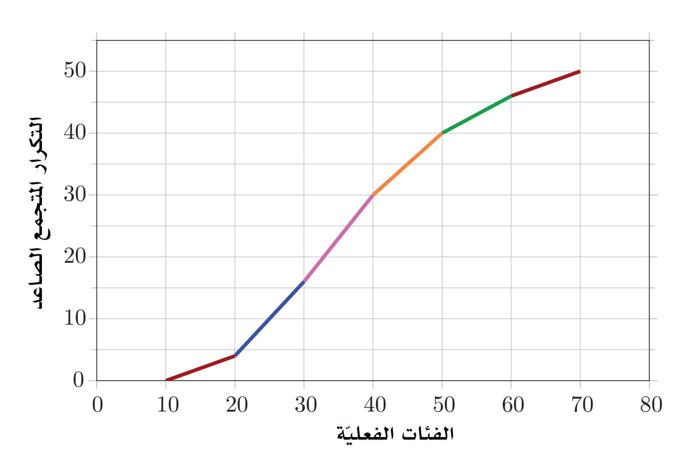
النقطة الأولى مسقطها على المحور الأفقي هو الحد الأعلى للفئة الفعليّة الأولى، ومسقطها على محور التكرارات يساوي قيمة التكرار المتجمِّع الصاعد للفئة الفعليّة الأولى.

النقطة الثانية مسقطها على المحور الأفقي هو الحد الأعلى للفئة الفعليّة الثانية، ومسقطها على محور التكرارات يساوي قيمة التكرار المتجمِّع الصاعد للفئة الفعليّة الثانية.

وهكذا دواليك حتى يتمّ تعيين النقطة الأخيرة التي مسقطها على المحور الأفقي هو الحد الأعلى للفئة الفعليّة الأخيرة، ومسقطها على محور التكرارات يساوي قيمة التكرار المتجمِّع الصاعد للفئة الفعليّة الأخيرة.

بعد ذلك نصِل بقطع مستقيمة بين النقاط التي تمَّ تعيينها آنفاً، وأخيراً نقوم بإغلاق بداية الضلّع الناتج إلى بداية الفئة الأولى فقط دون إغلاقه من اليمين (**أي أنَّه يغلق من طرف واحد فقط**).

فعلى سبيل المثال لو رجعنا إلى المثال (١-٤-١-١) فإنَّنا نجد أنَّ مضلَّع التكرار المتجمِّع الصاعد لبيانات المسافة المقطوعة لكل ليتر من الوقود له الشكل [1-7-c].



الشكل [1-7-c]

١-٤-٤- المنحنى التّكراري Frequency Curve

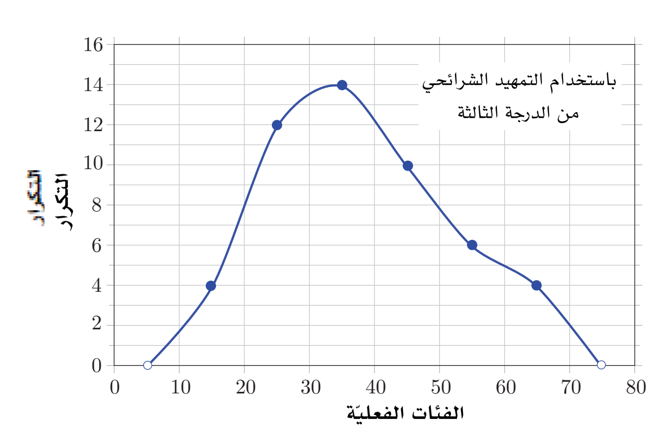
إنَّ هذه الطريقة تماثل طريقة العرض باستخدام المضلّعات التكرارية من حيث تعيين النّقاط التي سيمرّ منها الرّسم البياني، ولكنَّ الإغلاق من اليسار يكون إلى بداية الفئة الفعليّة الأولى، في حين يكون الإغلاق من اليمين إلى نهاية الفئة الفعليّة الأخيرة، ونحصل على هذا المنحني من خلال تمهيد الخط المنكسر ليصبح منحنياً وفقاً لإحدى الطرائق الآتية:

١- جعل الفترات التي تبنى عليها هذا الخطوط المنكسرة قصيرة جداً مع زيادة عددها بشكل كبير جداً، ولذلك فإنَّ هذه الطريقة نادرة الاستخدام في المجالات التّطبيقية.

٢- التمهيد اليدوي للخطوط المستقيمة لتصبح منحنية وملائمةً عند مرورها بالنقاط الممثّلة للبيانات، وهذه الطريقة قليلة الاستخدام أيضاً لأنَّها تحتاج إلى مهارة في الرسم.

٣- استخدام التمهيد الشرائحي Spline من الدرجة الثانية أو الثالثة، ومن ثمَّ النظر أيهما أنسب للعرض وأقل تناقضاً مع واقع البيانات، وهذه الطريقة تعطينا نتائج ممتازة لعمليّة التمهيد وهي متوفّرة في بعض البرامج الإحصائية، ومنها برنامج Curve Expert.

فعلى سبيل المثال لو رجعنا إلى المثال السابق (١-٤-١-١) نجد أنَّ المنحنى التكراري لبيانات جدول التوزيع التكراري المُعطى له الشكل الآتي.



الشكل [1-7-d]

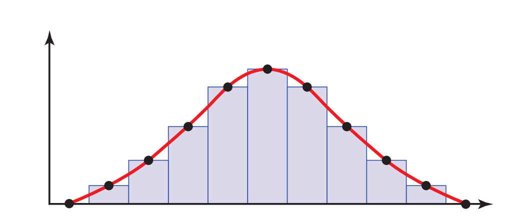
١ - ٥
أشكال التوزيعات التكرارية

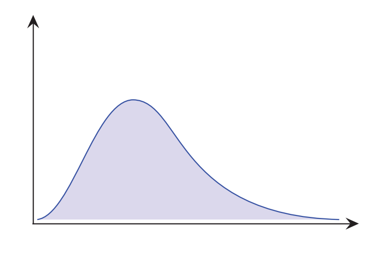
في الواقع توجد أشكال عديدة للتوزيعات التكرارية، وشكلها يتبع طبيعة انتشار البيانات ونزوعها نحو قيمة أو موضع ما، وهذه الأشكال تعطي دلالات مهمّة عن المؤثِّرات التي خضعت لها البيانات. في هذه الجزئية سوف لن نخوض بعيداً في هذه المسائل ولكن سنعرض بشكل موجز وبسيط أهم التوصيفات التي تذكر بها الأشكال البيانية للتوزيعات التكرارية، وسنبدأها بالفقرة الآتية.

١-٥-١- التوزيعات ذات المناويل Distributions that have Modes

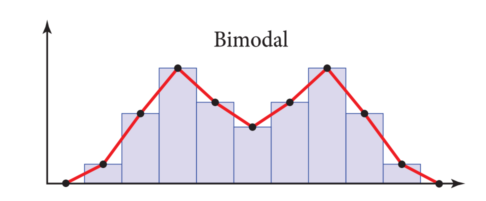
يمكن للمرء التمييز بين نوعين من التوزيعات هما:

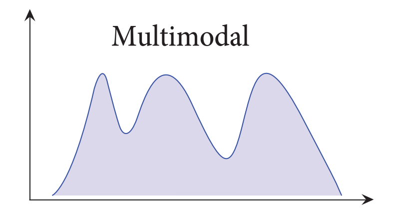
أ- توزيعات تملك قمّة (أو ذروة) واحدة، وهذه التوزيعات تُدعى توزيعات أحادية المنوال، والعرضين الآتيين يوضِّحان ذلك.





ب- توزيعات تملك أكثر من قمّة (أو ذروة)، وهذه التوزيعات تُدعى توزيعات متعددة المناويل.

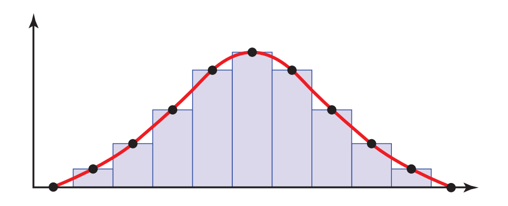


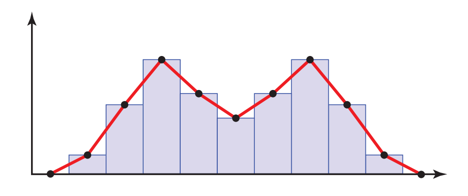


١-٥-٢- الالتواء والتناظر لتوزيع تكراري

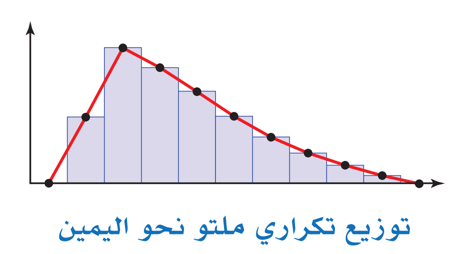
Skewness and Symmetry of Frequency Distributions

يُقال عن توزيع تكراري إنَّه متناظر إذا انطبق على نفسه تمام الانطباق لدى طيّه على محورٍ مارٍ من منتصف قاعدته، وفي حال عدم تحقّق الانطباق التام فإنَّه يُقال عن التوزيع التكراري إنَّه ملتوٍ، والعروض الآتية توضِّح لنا ذلك.

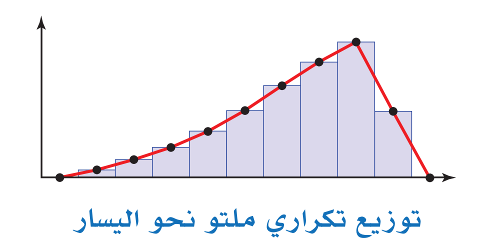




توزيعان تكراريان متناظران (أو متماثلان) Symmetric Frequency Distributions



توزيع تكراري ملتو نحو اليمين



توزيع تكراري ملتو نحو اليسار

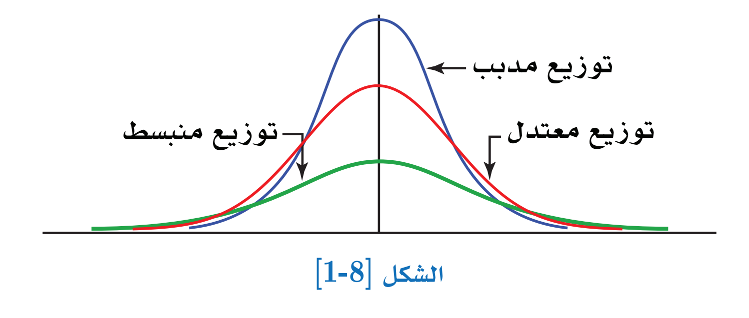
١-٥-٢- التدبب والتفلطح لتوزيع تكراري

Sputtering and Flattening of Frequency Distributions

تصنَّف التوزيعات التكرارية في ثلاثة أنواع (انظر الشكل الآتي [1-8]) هي:

أ- التوزيعات المدببة ب- التوزيعات المعتدلة ج- التوزيعات المنبسطة

علماً أنَّ مفهوم الاعتدال للتوزيعات بُنيَّ على مدى تقارب شكله من شكل دالّة الكثافة للتوزيع الطبيعي المعيَّاري (سنأتي على ذكر دالّة الكثافة للتوزيع الطبيعي المعيَّاري في الفصل الأخير)، وأمّا من أجل الحكم على تدبب أو تفلطح توزيع تكراري ما فإنَّنا نحتاج لمعيَّار مُحدّد لن نتطرق إليه هنا.



الشكل [1-8]

١-٥-٣- تفسير شكل التوزيع

في الواقع يمكن تقديم بعض التفسيرات لأشكال التوزيعات، ومنها:

١- إذا كان للتوزيع التكراري شكل متناظر معتدل فإنَّ ذلك يعني أنَّ البيانات تتوزع طبيعياً على وجه التقريب، وأنَّ الأخطاء المرتكبة لدى عمليّة القياس هي على الغالب أخطاء عشوائية (غير متعمَّدة).

٢- إذا كان للتوزيع التكراري عدَّة مناويل فإنَّ ذلك يدلِّل على وجود عدَّة أسباب فاعلة ومؤثِّرة في التجربة (أو المسألة) المولِّدة للبيانات، وعادة يكون عدد هذه الأسباب الفاعلة والمؤثِّرة مساوياً لعدد المناويل في التوزيع.

٣- إذا كان للتوزيع التكراري شكل ملتوٍ نحو اليمين فإنَّ ذلك يعني على الغالب تنفيذ عمليات فرز بشأن البيانات بحيث يستثنى منها البيانات التي تقل عن مقدار محدَّد تفرضه طبيعة الدراسة الإحصائية، وأمَّا إذا كان شكل التوزيع ملتوٍ نحو اليسار فإنَّ ذلك يعني على الغالب تنفيذ عمليات فرز بشأن البيانات بحيث يستثنى منها البيانات التي تزيد عن مقدار محدَّد تفرضه طبيعة الدراسة الإحصائية.

٤- إذا كان شكل التوزيع التكراري منبسطاً فإنَّ ذلك يعني على الغالب تنفيذ عمليات فرزٍ بشأن البيانات بحيث يُستثنى منها البيانات التي تقل وتزيد عن قيمة محدَّدة تفرضها طبيعة الدراسة الإحصائية، وكذلك يدلِّل على أنَّ البيانات تتبعثر بعيداً عن قيمة المتوسط لهذه البيانات وذلك تبعاً لمقدار تفرطح (أو تفلطح) هذا التوزيع.



تمارين



١- وضح الفرق بين العيِّنة الإحصائية والمجتمع.

٢- لماذا نحتاج إلى تبويب (تفريغ) البيانات في جداول تكرارية، وضح ذلك بالتفصيل.

٣- صنف المتغيِّرات الآتية من حيث كونها متغيِّرات كميَّة أو متغيِّرات نوعيّة:

أ- عدد الطلاب في جامعة الملك فهد للبترول والمعادن.

ب- فصيلة الدم لعيِّنة مكونة من 50 طالباً.

ج- أنواع التمور في مزرعة معيِّنة.

د- جنسية مجموعة من الأشخاص يعيشون بالمملكة العربية السعودية.

هـ- تقديرات مجموعة من الطلاب في مقرَّر الفيزياء.

و- عدد الكيلومترات التي يقطعها الطلاب للذهاب إلى جامعة الملك سعود.

ز- أعمار الطلاب في جامعة الملك عبد العزيز.

ح- عدد الوجبات الغذائية المقدَّمة في مطعم معيَّن في يوم ما.

ك- مستوى الخدمة المقدَّم من فندق معيَّن في مكة المكرمة.

٤- حدد فيما إذا كانت العبارات الآتية صحيحة أم خاطئة:

أ- العمر يعدّ مثالاً على متغيِّر نوعي.

ب- عدد الكليات في جامعة الملك سعود يعدّ متغيِّراً متقطِّعاً.

ج-عدد الدقائق التي يقطعها الطالب للوصول إلى جامعة الإمام محمد بن سعود الإسلامية تعدّ متغيِّراً متّصلاً.

د- لا يوجد علاقة بين المجتمع الإحصائي والعيِّنة الإحصائية.

ه- المتغيِّر الذي يمكن تمثيل قيمه على شكل فترات يسمى متغيِّر نوعي.

٥- عدد أنواع البيانات الإحصائية، ومن ثمَّ اذكر مثالاً على كل نوع منها.

٦- ليكن لدينا جدول التوزيع التكراري الآتي:

رقم الفئة	الفئات الفعليّة	مراكز الفئات	التكرار	التكرار المتجمع الصاعد
1	10 – 14	12	2	2
2	14 –18	16	4	6
3	18 – 22	20	8	14
4	22 – 26	24	16	30
5	26 – 30	28	10	40
Total			**40**	

والمطلوب ما يلي:

أ- رسم المدرَّج التكراري لبيانات هذا الجدول.

ب- رسم المضلَّع التكراري لبيانات هذا الجدول.

ج- رسم مضلَّع التكرار المتجمِّع الصاعد لبيانات هذا الجدول.

د- رسم المنحني التكراري لبيانات هذا الجدول مستخدماً التمهيد باليد.

٧- ليكن لدينا جدول التوزيع التكراري الآتي:

الفئات الفعليّة	6-9	9-12	12-15	15-18	18-21	21-24	المجموع
التكرار	4		16			4	**50**
التكرار النسبي		0.12					
التكرار المئوي				24 %			
التكرار المتجمع الصاعد					46		

والمطلوب إكمال مُعطيات هذا الجدول.

٨- أخذت عيِّنة مكونة من 100 شخص وتمَّ تصنيفهم حسب عدد مرات الذّهاب لنادي رياضي معيّن خلال شهرٍ، فكانت النتائج كما في جدول التوزيع التكراري الآتي:

عدد مرات الذهاب إلى النادي	عدد الأشخاص
0 – 3	23
4 – 7	40
8 – 11	28
12 – 14	6
16 – 18	3
Total	**100**

والمطلوب ما يلي:

أ- عيّن مراكز الفئات.

ب- ما هو طول الفئة لهذا الجدول التكراري؟

ج- قم ببناء الجدول التكراري النسبي والتكراري المئوي؟

د- ما هي نسبة الأشخاص الذين يذهبون إلى النّادي سبع مرّات على الأكثر خلال الشّهر؟

٩- أخذت عيِّنة عشوائية من إحدى المدارس مكوَّنة من 50 طالباً، وقُرأت قيمة الوزن لكل واحدٍ منهم (مقدَّرةً بالكيلوجرام)، فكانت لدينا النتائج الآتية:

30	34	27	23	33	33	26	25	24	28
21	26	31	22	27	33	27	23	28	21
31	35	34	22	26	25	23	35	31	27
30	34	27	22	27	33	23	35	31	27
34	22	26	33	33	26	27	23	22	27

والمطلوب ما يلي:

أ- صبُّ هذه البيانات في جدول توزيع تكراري.

ب- رسم المدرَّج التكراري، المضلَّع التكراري ومضلَّع التكرار المتجمّع الصاعد لبيانات هذا الجدول.

ج- باستخدام مضلَّع التكرار المتجمّع الصاعد للبيانات قدِّر التكرار المتجمّع الصاعد للقيمة 25.5.

د- حدِّد شكل التوزيع من حيث كونه متماثلاً أم لا.

١٠- البيانات الآتية تُمثِّل عيّنة نتائج استبيانات مُصنَّفة في ثلاثة أنواع A، B وC كما يلي:

B	A	C	C	A	C	B	A	C	B
C	C	C	B	C	B	C	B	C	C
B	C	C	C	A	B	C	C	B	C

والمطلوب ما يلي:

أ- صبُّ هذه البيانات في جدول تكراري تظهر فيه التكرارات، التكرارات النسبيّة والتكرارات المئوية.

ب- تقديم العرض الشرائطي (**بالأعمدة**) لهذه البيانات.

١١- أُخِذت عيّنة مكوَّنة من 80 شخصاً، وسُئل كلُّ واحدٍ منهم عن آخر مؤهلٍ علميٍّ حصل عليه فكانت النّتائج موضَّحة كما في الجدول الآتي:

مستوى التعليم	ابتدائي	متوسط	ثانوي	جامعي	دراسات عليا	المجموع
عدد الأشخاص	10	16	18	30	6	**80**

والمطلوب تمثيل بيانات هذا الجدول باستخدام القطاعات الدّائريّة.

١٢- من أحد الحقول الزّراعيّة تم أخذ عيِّنة مكونة من 40 زهرة وتمّ تصنيف هذه الزّهور بحسب اللون، فكانت لدينا النتائج الآتية:

أصفر	أحمر	أبيض	أصفر	أصفر	أحمر	أبيض	بنفسجي
أصفر	أبيض	أحمر	بنفسجي	أصفر	أصفر	أحمر	بنفسجي
أحمر	أصفر	أبيض	أصفر	بنفسجي	أحمر	أبيض	أصفر
أصفر	أحمر	بنفسجي	أحمر	أصفر	أحمر	أبيض	أبيض
أصفر	أبيض	أصفر	بنفسجي	أصفر	أحمر	أصفر	بنفسجي

والمطلوب ما يلي:

أ- صبُّ هذه البيانات في جدولٍ تكراريٍّ يظهر فيه التكراري النسبي والمئوي لهذه البيانات.

ب- تمثيل هذه البيانات بالعرض الشّرائطي (**بالأشرطة الأفقية**).

ج- تمثيل هذه البيانات بالقرص الدّائري وموضِّحاً قيم زوايا القطاعات الدائرية الناتجة.

د- ما هي نسبة الزّهور الحمراء؟

١٣- الجدول الآتي يوضّح وقت الانتظار (**بالدقائق**) في طوارئ أحد المستشفيات لعيِّنةٍ مكوَّنةٍ من 90 مريضاً.

عدد المرضى	وقت الانتظار
5	0 – 6
15	7 – 13
20	4 – 20
15	21 – 27
5	28 – 34
60	**Total**

مستخدماً الجدول السّابق اختر الإجابة الصحيحة لكلٍ من العبارات الآتية:

أ- عدد الفئات في هذا الجدول هو: 4 ، 5 ، 6 أو 7.

ب- سعة (أو طول) الفئة لبيانات الجدول هي: 8 ، 5 ، 6 أو 10.

ج- مركز الفئة الثالثة هو: 15.5 ، 16.5 ، 17.5 أو 19.5.

د- الحد الأدنى العملي للفئة الثانية هو: 6 ، 6.5 ، 7 أو 7.5.

ه- الحد الأعلى للفئة الفعليّة الثالثة هو: 19.5 ، 20 ، 20.5 أو 21.

و- حجم العيِّنة هو: 48 ، 58 ، 60 أو 65.

ز- التكرار النسبي للفئة الرابعة هو: 0.20 ، 0.22 ، 0.25 أو 0.27.

ح- التكرار المئوي للفئة الأخيرة هو: 7.8 ، 8.3 ، 8.5 أو 9.2.

ك- التكرار المتجمّع الصاعد للفئة الأولى هو: 0 ، 5 ، 15 أو 20.

ل- توزيع هذه البيانات: ملتوٍ نحو اليمين، ملتوٍ نحو اليسار أم متناظر.

م- توزيع هذه البيانات: أحادي المنوال، ثنائي المنوال أم لا منوال له.

١٤- الجدول الآتي يعطينا مبيعات شركتين A وB (مقدَّرة بالملايين) خلال خمسة أشهر.

الشهر	مبيعات الشركة A	مبيعات الشركة B
يناير	5.6	3.8
فبراير	4.7	3.3
مارس	3.9	2.8
إبريل	4.5	3.5
مايو	6.1	4.2

والمطلوب تمثيل هذه البيانات باستخدام الشرائط العمودية المزدوجة (الأعمدة البيانية المزدوجة).

١٥- الجدول الآتي يمثّل مبيعات أحد محلات القهوة بالرّيال خلال فترتين من اليوم ولمدّة أسبوع كامل.

اليوم	مبيعات الفترة الأولى	مبيعات الفترة الثانية	المبيعات الكلية
السبت	530	150	**680**
الأحد	670	270	**940**
الإثنين	550	160	**710**
الثلاثاء	630	250	**880**
الأربعاء	620	230	**850**
الخميس	740	490	**1230**
الجمعة	850	530	**1380**

والمطلوب ما يلي:

أ- مثّل البيانات السابقة باستخدام الأعمدة البيانية المزدوجة.

ب- مثّل البيانات السابقة باستخدام الأعمدة البيانية المجزأة.

ج- أيٌّ من التمثيلين السابقين أفضل، ولماذا؟

١٦- البيانات الآتية تمثّل عدد الكيلومترات التي يقطعها أربعون مهندساً للوصول إلى مقر عملهم.

5	3	10	20	25	11	13	7	12	31
19	10	12	17	18	11	32	17	16	2
7	9	7	8	3	5	12	15	18	3
12	14	2	9	6	15	15	7	6	12

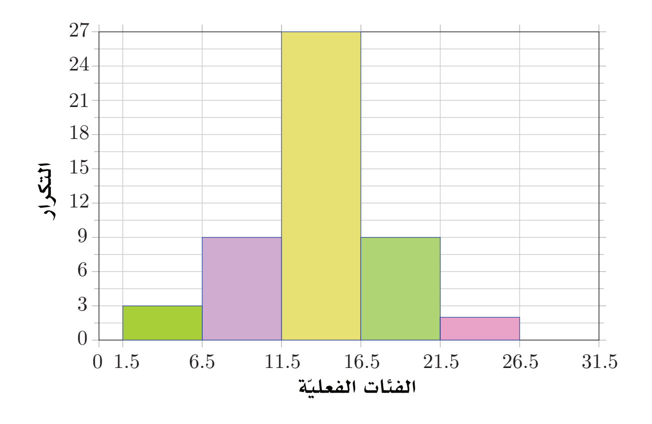
والمطلوب ما يلي:

أ- صبُّ هذه البيانات في جدولٍ توزيع تكراري تكون فيه الفئة الفعليّة الأولى هي 5 — 0 .

ب- كم عدد المهندسين الذين يقطعون مسافة 20 كيلو على الأكثر للوصول إلى مقرِّ عملهم؟

ج- كم عدد المهندسين الذين يقطعون مسافة 15 كيلو على الأقل للوصول إلى مقرِّ عملهم؟

١٧- لتكن لدينا بيانات مقدّمة من خلال العرض البياني الآتي:



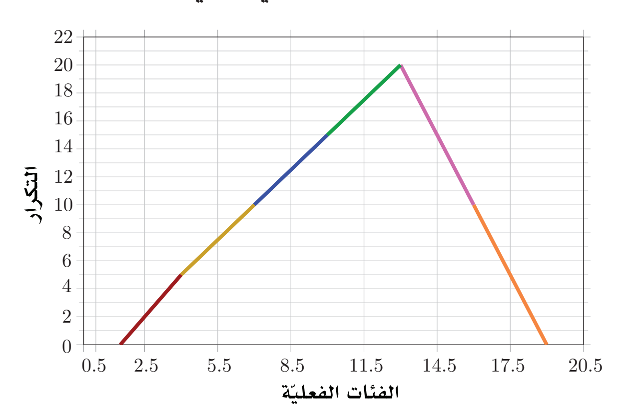
والمطلوب ما يلي:

أ- قم ببناء جدول التوزيع التكراري لهذه البيانات.

ب- ارسم المضلَّع التكراري ومضلَّع التكرار المتجمِّع الصاعد لهذه البيانات.

ج- هل يوحي هذا الشكل إلى أنَّ توزيع البيانات متناظر تماماً، ولماذا؟

١٨- لتكن لدينا بيانات مقدَّمة من خلال العرض البياني الآتي:



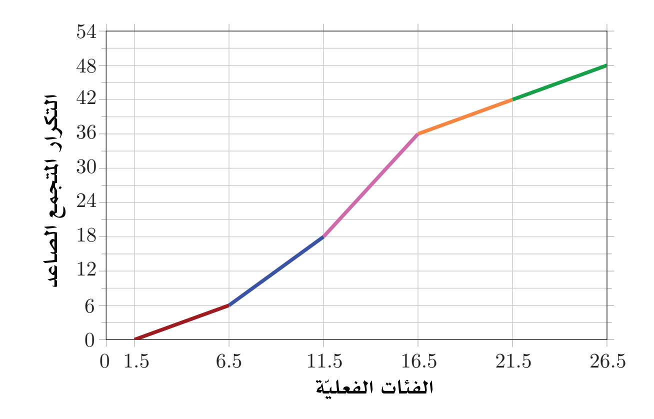
والمطلوب ما يلي:

أ- قم ببناء جدول التوزيع التكراري لهذه البيانات.

ب- ارسم المدرَّج التكراري ومضلَّع التكرار المتجمِّع الصاعد لهذه البيانات.

ج- إلى أية جهةٍ يلتوي فيها شكل توزيع البيانات؟

١٩- لتكن لدينا بيانات مقدَّمة من خلال العرض البياني الآتي:



والمطلوب ما يلي:

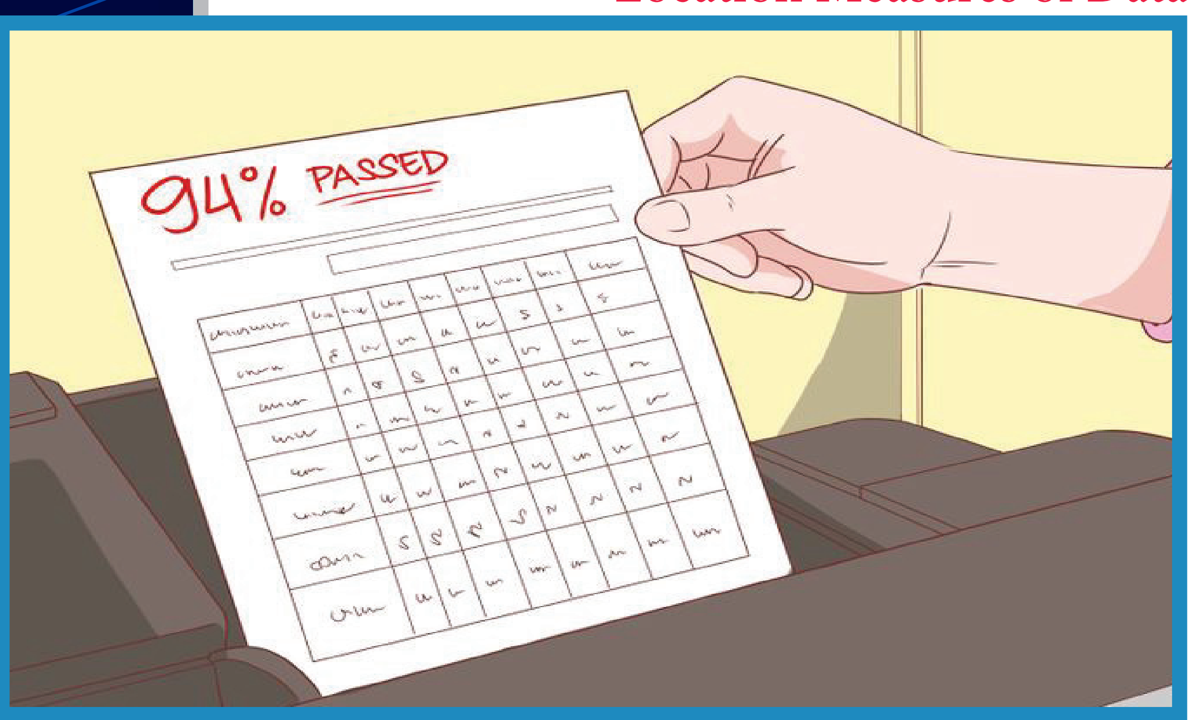
أ- قم ببناء جدول التوزيع التكراري لهذه البيانات.

ب- ارسم المدرَّج التكراري والمضلَّع التكراري لهذه البيانات.

ج- هل يوحي لك هذا الشكل إلى وجود التواء في توزيع البيانات، ولماذا؟

الفصل الثاني

مقاييس الموضع للبيانات
Location Measures of Data



المقدمة:

لقد لاحظنا أنَّ طرائق العرض البيانية لها أهمية خاصة عند تقديم المعلومات الإحصائية فهي تُعطي وصفاً عاماً وسريعاً للبيانات الإحصائية، إلاَّ أنَّ فوائدها الاستقرائية تبقى قليلة، فإذا نظرنا إلى المضلَّع التكراري لمجموعة بيانات عينة فإنَّه يعطينا تصوراً عن شكل المضلَّع التكراري لهذه البيانات، واستقراؤنا له يقف عند الفرض أنَّه يوجد تشابه ما بين المضلَّع المثِّل لهذه البيانات والمضلَّع الممثِّل لمجتمع البيانات التي أُخذت منها هذه العينة، ولذلك كان لا بدَّ من تقديم معايير عدديّة تحدَّد لنا وبدّقة سلوك البيانات. من المعايير التي تقوم بهذه المهمّة ما يُعرَّف باسم «مقاييس الموضع» وعلى وجه الخصوص مقاييس النزعة المركزيّة.

- ٢ - ١ - مقاييس النزعة المركزيّة
- ٢ - ٢ - الربيعيَّات
- ٢ - ٣ - المئينات
- ٢ - ٤ - الأعداد الخمسة والتمثيل الصندوقي للبيانات

٢ - ١
مقاييس النزعة المركزية

في الواقع لكلِّ مجتمع مَعَالم (أو وسطاء) Parameters تميّزه، وهذه المَعالم تكون عادة مجهولةً كلياً أو جزئياً، ولكي نتمكن من التعامل مع المجتمع الإحصائي يجب علينا تقدير مَعَالمه المجهولة. إنَّ الوسائط التي تقوم بهذه المهمّة تُدعى الإحصاءات (مفردها إحصاء Statistic)، وهذه الإحصاءات تتعامل مع بيانات العينات مباشرة. لذلك لا بدّ لنا أن نتعرّف أولاً على مفهومي المَعْلَمة والإحصاء.

المَعْلَمة: إنَّ أية قيمة عددية تُميِّز المجتمع تُدعى مَعْلَمةً (وفي بعض المراجع يُقال عنها وسيط، ومنها على سبيل المثال: المتوسط والانحراف المعياري للمجتمع- سنأتي على ذكرهما لاحقاً-)، وهذه القيمة تكون على الغالب مجهولة ويجب تقديرها من بيانات عينة تسحب من هذا المجتمع.

الإحصاء: إنَّ أية قيمةٍ عدديةٍ تُميِّز العينة تُدعى إحصاءً (ومنها على سبيل المثال: المتوسط والانحراف المعياري للعينة)، وتحسب هذه القيمة من بيانات العينة، ومن ثمَّ فهي قيمة يمكن الحصول عليها بالحساب، وبالتالي يُنظر إليها على أنَّها قيمة معلومة وتستخدم كقيمةٍ تقريبيةٍ للمَعْلَمة.

الآن بالعودة إلى الحوار السّابق فإنَّ المشكلة التي سنقف عندها هي كيفية قياس درجة الاختلاف بين قيمة مَعْلَمة المجتمع وما يقابلها من قيمة مَعْلَمة العيِّنة (قيمة الإحصاء). لهذا السبب كان لا بدَّ من وجود مقاييس كميَّة تُمكّننا من تقدير مَعَالم المجتمع الإحصائي.

فيما يلي سنوجّه اهتمامنا على استخراج قيمة أو أكثر من مجموعة البيانات للاستدلال من خلالها على حقائق الظاهرة التي تمثّلها مجموعة البيانات ككلّ، وهذا يتطلب منّا الآتي:

أ- البحث عن قيمة عددية (أو قيم عددية) تُمثِّل مركز جذب لهذه البيانات، بمعنى آخر، تبدو البيانات كلّها أو بعضها ينزع نحو هذه القيمة (أو نحو هذه القيم)، وهذا المبحث سيكون محور هذا الفصل من الكتاب.

ب- البحث عن قيمة عددية توضِّح مدى تبعثر قيم البيانات عن تلك القيمة التي تنزع إليها البيانات، وهذا المبحث سيكون محور الفصل الثالث من هذا الكتاب.

ج- البحث عن قيم عددية تدل على شكل التوزيع من حيث الالتواء والتفلطح، بمعنى أنَّ هذه القيم تحدِّد لنا وبدقّة إن كان التوزيع ملتوياً أم لا، وكذلك إن كان التوزيع مفلطحاً، مدبباً أم معتدلاً، وهذا المبحث لن نتناوله في كتابنا هذا.

كما يوجد أنوَّاع أخرى من المقاييس التي تساعدنا في عملية الاستقراء لنتائج البيانات، إلا أنَّها أقل أهمية من المقاييس السابقة.

لقد تحدثنا قبل قليل حول ضرورة البحث عن القيمة العـددية (أو القيم العددية) التي تنزع نحوها البيانات التي قيد الدراسة. إنَّ هذه القيمة (أو القيم) تندرج تحت مفهوم مقاييس النزعة المركزيّة والتي تقدّمها لنا الفقرة الآتية.

٢-١-١- تعريف (مقياس النزعة المركزيّة)

لتكن لدينا مجموعة بيانات مُعطاة. عندئذٍ كل قيمة عددية تنزع إليها البيانات (تُبدو كمركز جذب للبيانات) كلياً أو جزئياً تُدعى **مقياساً للنزعة المركزيّة**.

من التعريف السابق نلاحظ أن القيمة العددية التي تمثّل مقياساً للنزعة المركزيّة تُعبّر في الواقع عن موضع تمركُز توزيع البيانات، ومن المقاييس التي تهتمّ بهذه الدراسة **المتوسّط** (أو الوسط الحسابي)، **المتوسّط الهندسي**، **المتوسّط التوافقي**، **الوسيط والمنوال**، ولكن في كتابنا هذا سنقوم بالتركيز على استخدام **المتوسّط، الوسيط والمنوال** فقط، والتي سنقوم بتقديمها تباعاً.

٢-١-٢- تعريف (المتوسّط) (Mean)

من أجل تعريف المتوسّط لمجموعة بيانات مُعطاة سنناقش الحالات الثلاث الآتية:

١- إذا كانت البيانات المُعطاة **خام** من قبيل x_1، x_2، ... و x_n (قد لا تكون مختلفة بعضها عن البعض الآخر)، فعندئذٍ يُعرَّف المتوسط لهذه البيانات على أنَّه قيمة عددية (يُرمز لها بـ $\bar{x}$) تحسب بوساطة العلاقة الآتية:

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n} \sum_{i=1}^{n} x_i \qquad [2\text{–}1\text{-a}]$$

٢- إذا كانت البيانات المُعطاة **مفردة** (خام) من قبيل x_1، x_2، ... و x_k ولها أوزان w_1، w_2، ... و w_k، فعندئذٍ يُعرَّف المتوسط لهذه البيانات على أنَّه قيمة عددية (يُرمز لها بـ $\bar{x}_w$) تحسب بوساطة العلاقة الآتية:

$$\bar{x}_w = \frac{w_1 \cdot x_1 + w_2 \cdot x_2 + \cdots + w_k \cdot x_k}{w_1 + w_2 + \cdots + w_k} = \frac{\sum_{i=1}^{n} w_i \cdot x_i}{\sum_{i=1}^{n} w_i} \qquad [2\text{–}1\text{-b}]$$

إنَّ القيمة العددية $\bar{x}_w$ تُدعى المتوسط الموزون للبيانات x_1، x_2، ... و x_k.

٣- إذا كانت البيانات المُعطاة **مجمَّعةً** في جدول توزيع تكراري بـ k فئة كما في الجدول الآتي:

الجدول [2-1]

رقم الفئة	الحدود الفعليّة للفئة	مركز الفئة	تكرار الفئة	التكرار المتجمّع الصاعد للفئة
1	$a_0 \to a_1$	x_1	f_1	$F_1 = f_1$
2	$a_1 \to a_2$	x_2	f_2	$F_2 = f_1 + f_2$
⋮	⋮ ⋮ ⋮	⋮	⋮	⋮ ⋮ ⋮ ⋮
k	$a_{k-1} \to a_k$	x_k	f_k	$F_k = f_1 + f_2 + \ldots + f_k$
Total	----------------	--------	$\sum_{i=1}^{k} f_i$	المجموع

فعندئذٍ يحسب المتوسط لبيانات هذا الجدول من خلال العلاقة الآتية:

$$\bar{x} = \frac{1}{\sum_{i=1}^{k} f_i} \sum_{i=1}^{k} f_i \cdot x_i \qquad [2\text{–}1\text{-c}]$$

إنَّ العلاقة الأخيرة توضِّح لنا أنَّه يمكن تمثيل كل فئة من فئات الجدول التكراري بمركزها كما ذكرنا ذلك في الفصل السابق، وأنَّ لهذا المركز وزناً قدره يساوي إلى تكرار هذه الفئة.

٢-١-٢-١- أمثلة

١- قام تاجر بشراء أربعة من الإبل كلّ على انفراد بثمنٍ قدره 5950، 6950، 7950 و6250 ريال. وأراد بيعها دفعةً واحدة بسعر الحبّة الواحدة، فما هو الحد الأدنى للمبلغ الذي يجب أن يطلبه من الشاري في الحبّة الواحدة حتّى لا يقع في أية خسارة؟

الجواب: حتى لا يقع التاجر في أية خسارة يجب ألاّ يقل المبلغ الذي سيطلبه في الحبّة الواحدة عن قيمة متوسط الثمن الذي دفعه في الإبل الأربعة، ومن ثمَّ بتطبيق العلاقة [2–1-a] يكون لدينا:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i = \frac{5950 + 6950 + 7950 + 6250}{4} = \frac{27100}{4} = 6775$$

إذن عليه أن يطلب 6775 ريالاً على الأقل ثمناً لكلّ رأس من الإبل حتى لا يقع في أية خسارة.

٢- قام طالب بحساب الزمن (مقدّراً بالدقيقة) الذي يستغرقه الطريق من منزله إلى عمادة السّنة الأولى المشتركة، وذلك على مدى 12 يوماً، فحصل على القيم الآتية:

37	33	42	27	39	25	39	37	28	36	28	32

فما هو متوسط الزمن اللازم لتنقل الطالب من منزله إلى العمادة؟

الجواب: بتطبيق العلاقة [2–1-a] نجد أنّ متوسط الزمن الـلازم لتنقـل الطالـب مـن منزلـه إلى العمادة يساوي:

$$\bar{x} = \frac{37+33+42+27+39+25+39+37+28+36+28+32}{12}$$

$$= \frac{403}{12} = 33.58$$

إذن يحتاج الطالب بالمتوسط إلى 33.58 دقيقة للتنقل من منزله إلى عمادة السّنة الأولى المشتركة.

٣- إذا كان لدى طالب خمس مقرَّرات دراسيّة، ولكل مقرَّر من هذه المقرَّرات وزن خاص به مقدَّم في الجدول الآتي، وبفرض أنَّ الطالب قد نجح في هذه المقرَّرات الخمس وكانت درجاته في كل مقرَّر كما هو مدوَّن في الجدول الآتي أيضاً، فما هو متوسط درجات هذا الطالب في المقرَّرات الخمس هذه؟

الجدول [2-2]

المقرَّر	لغة إنكليزية	مهارات الاتصال	تقن	إحصاء	تطوير الذات
الوزن w	3	2	2	1	1
الدرجة النهائية للطالب	87	92	80	93	98

الجواب: بناءً على المعطيات المقدَّمة نجد أنَّ متوسط درجات هذا الطالـب في المقـرَّرات الخمـس هو المتوسط الموزون لهذه الدرجات، ومنه بتطبيق العلاقة [2–1-b] يكون لدينا:

$$\bar{x}_w = \frac{1}{\sum_{i=1}^{k} w_i} \sum_{i=1}^{k} w_i \cdot x_i = \frac{(3\times 87)+(2\times 92)+(2\times 80)+(1\times 93)+(1\times 98)}{9}$$

$$= \frac{796}{9} = 88.\bar{4}$$

٤- ليكن لدينا بيانات إحصائية مُعطاة من خلال جدول التوزيع التكراري الآتي الذي يعرض عدد الزبائن الذين اشتروا بضائع (مقدَّرة بالريال) من سوق تجاري في يوم معيَّن، فما هو متوسط مبلغ المبيعات لكل زبون؟

الجدول [2-3-a]

رقم الفئة	(نطاق المبلغ الذي اشترى به الزبون) الحدود الفعليّة للفئة	(عدد الزبائن) تكرار الفئة
1	$0 \rightarrow 100$	125
2	$100 \rightarrow 200$	48
3	$200 \rightarrow 300$	32
4	$300 \rightarrow 400$	25
5	$400 \rightarrow 500$	27
Total	------------------------------------	**257**

الجواب: من المعلوم أنَّ قيمة المتوسط لبيانات جدول ذو فئات يُعطى بالعلاقة [2–1-c]، ومن ثمَّ يتوجب علينا تعيين مراكز الفئات لإتمام عملية الحساب حيث لدينا:

الجدول [2-3-b]

رقم الفئة	5	4	3	2	1
الحدود الفعليّة للفئة	400 → 500	300 → 400	200 → 300	100 → 200	0 → 100
مركز الفئة	450	350	250	150	50

ومنه يكون متوسط مبلغ المبيعات لكل زبون يساوي:

$$\bar{x} = \frac{1}{\sum_{i=1}^{k} f_i} \sum_{i=1}^{k} f_i \cdot x_i$$

$$= \frac{(125 \times 50) + (48 \times 150) + (32 \times 250) + (25 \times 350) + (27 \times 450)}{257}$$

$$= \frac{42350}{257} = 164.79$$

إذن متوسط مبلغ المبيعات لكل زبون هو 164.79 ريالاً.

٥- بالرجوع إلى المثال (١-٣-٤) (عدد الأميال المقطوعة لكل لتر من البنزين لأربعين سيارة حديثة) نجد بتطبيق العلاقة [2–1-c] أنَّ لبيانات جدول التوزيع التكراري المُعطى متوسط يساوي:

$$\bar{x} = \frac{1}{\sum_{i=1}^{k} f_i} \sum_{i=1}^{k} f_i \cdot x_i$$

$$= \frac{(8 \times 12.5) + (10 \times 14.5) + (12 \times 16.5) + (5 \times 18.5) + (5 \times 20.5)}{40}$$

$$= \frac{638}{40} = 15.95$$

٦- لتكن لدينا مجموعة من البيانات الخام الآتية التي تمثِّل الطول (مقدَّراً بالسنتيمتر) لخمسين طالباً.

175	190	175	172	160	182	155	143	179	160
177	184	179	171	183	165	184	168	160	182
190	168	182	168	157	153	182	175	159	169
149	168	173	162	183	162	144	191	178	162
199	140	185	176	169	166	151	167	160	197

فنجد أنَّ متوسط الطول لهؤلاء الطلاب يساوي:

$$\bar{x} = \frac{175 + 190 + 175 + \dots + 167 + 160 + 197}{50} = \frac{8529}{50} = 170.58$$

وبصبِّ هذه البيانات في جدول توزيع تكراري ذو فئات فعليّة بسعة تساوي 10 نجد أنَّ عدد الفئات المطلوبة لبناء جدول التوزيع التكراري يساوي $k = 6$ فئات، ومن ثمَّ يكون لجدول التوزيع التكراري العرض الآتي.

الجدول [2-4]

رقم الفئة	الحدود الفعليّة للفئة	مركز الفئة	تكرار الفئة	التكرار المتجمِّع الصاعد للفئة
1	140 → 150	145	4	4
2	150 → 160	155	5	9
3	160 → 170	165	16	25
4	170 → 180	175	11	36
5	180 → 190	185	10	46
6	190 → 200	195	4	50
Total	------------	--------	50	-----------------

فنجد أنَّ متوسط الطول لهؤلاء الطلاب بعد صبِّها في جدول التوزيع التكراري يساوي:

$$\bar{x} = \frac{1}{\sum_{i=1}^{k} f_i} \sum_{i=1}^{k} f_i \cdot x_i = \frac{(4 \times 145) + (5 \times 155) + \cdots + (4 \times 195)}{50} = \frac{8550}{50} = 171$$

لاحظ هنا أنَّ الاختلاف في قيمة المتوسط بين البيانات الخام والمجمَّعة مردّه عدم الأخذ بالحسبان جميع القيم والاكتفاء بقيم ممثِّلة للفئات (التي هي مراكز الفئات) فقط.

٢-١-٢-٢- ملاحظة

إنّ المجموع الجبري للفروق $x_1 - \bar{x}$ ، $x_2 - \bar{x}$ ، ... و $x_n - \bar{x}$ يساوي الصفر دوماً، وهذا يعني أنَّ المجموع الجبري لانحرافات قيم البيانات عن متوسطها معدوم دوماً، وذلك لأنَّ:

$$\sum_{i=1}^{n} \left(x_i - \bar{x}\right) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \bar{x} = \sum_{i=1}^{n} x_i - n \cdot \bar{x} = n \cdot \bar{x} - n \cdot \bar{x} = 0$$

٢-١-٢-٣- الميزات الإيجابية للمتوسط

إنَّ المتوسط أكثر مقاييس النزعة المركزيّة استخداماً وذلك يعود لسهولة حسابه وتعريفه بعلاقة رياضياتية بسيطة، ومن أهم الميزات الإيجابية التي يتمتَّع بها هي:

١- يأخذ بالحسبان جميع القياسات التي تخضع للدراسة والبحث، ولذلك يُفضَّل استخدامه عندما يكون الاهتمام منصباً على القيمة العددية التي تأخذ جميع القياسات بالحسبان وليس الحصول على قيمة نموذجية مُمثِّلة لها فقط.

٢- يخضع للعمليات الجبرية المعروفة، فلو كانت a و b عددين حقيقيين مع $a \neq 0$، فعندئذٍ:

أ- المتوسط للقيم $a \cdot x_1$، $a \cdot x_2$، ... و $a \cdot x_n$ يساوي $a \cdot \overline{x}$،

ب- المتوسط للقيم $x_1 + b$، $x_2 + b$، ... و $x_n + b$ يساوي $\overline{x} + b$،

ومن ثمَّ ينتج لدينا أنَّ المتوسط للقيم $a \cdot x_1 + b$، $a \cdot x_2 + b$، ... و $a \cdot x_n + b$ يساوي $a \cdot \overline{x} + b$.

٣- إذا عُلِمت قيمته فإنَّه يمكن حساب مجموع البيانات إذا كان عددها معلوماً، أو حساب عدد البيانات إذا كان المجموع لها معلوماً.

٤- من إحدى ميزاته الجميلة أيضاً، هي أنَّه إذا اضطررنا إلى تمديد حجم البيانات، فإنَّه بإضافة قيم تساوي قيمة متوسط البيانات الأصل لا تتغيَّر قيمة هذا المتوسط.

٥- يفضّل استخدام المتوسط عندما يكون التوزيع متماثلاً (أو متناظراً) على وجه التقريب، وكذلك عندما تكون جميع البيانات معلومة ولا يوجد فيها بيانات مفقودة.

٢-١-٢-٤- سلبيات المتوسط

١- إذا فقدت إحدى أو بعض قيم البيانات فإنَّ المتوسط يصبح عديم التطبيق (حتى إذا عُلم ترتيبها بين البيانات).

٢- تأثُّره بالقيم المتطرفة Extreme Values (أو القيم المنعزلة Outlier Values وسنأتي على تعيين هذه القيم لاحقاً في هذا الفصل)، فعلى سبيل المثال لو قامت مجموعة من التلاميذ بجمع تبرعات من أجل عمل خيري، وبفرض أنَّ هذه التبرعات كانت على النحو الآتي (مقدّرةً بالريال):

قيمة متطرفة

4 5 45 3 4 5 6 7 4 7

فعندئذٍ نجد أنَّ كل تبرعات التلاميذ كانت أقل من 8 باستثناء أحدهم قد تبرع بـ 45 ريالاً، فلو استثنينا هذه القيمة الأخيرة لوجدنا أنَّ متوسط تبرعات التلاميذ التسعة هو 5 ريالات فقط، ولكن عندما تبرع الطالب العاشر بـ 45 ريالاً أصبحت قيمة المتوسط 9 ريال، وكأنَّما قامت القيمة المتطرفة (45) بسحب قيمة المتوسط نحوها ممّا أظهر لنا نتيجة لا تُعبِّر عن واقع التبرعات للتلاميذ.

إذاً، فإذا وجدت قيم متطرفة فإنَّنا على الأقلّ غير قادرين على إعطاء قيمة مقنعة ومقبولة لقيمة المتوسط، ولذلك فمن غير المرغوب فيه استخدام المتوسط كمقياس للنزعة المركزيّة في هذه الحالة. كذلك إذا فقدت إحدى أو بعض البيانات فإنَّنا سنقف عاجزين عن حساب قيمة المتوسط للبيانات المُعطاة. لذلك كان من الضرورة البحث عن طرائق أخرى تقوم بهذه المهمّة. إنَّ المفهوم الآتي يقدِّم لنا حلاً جزئياً للمشكلة السابق ذكرها، ويعدّ من المفاهيم البديلة للمتوسط في حال عدم إمكانية استخدامه أو عدم القبول بنتيجته.

٢-١-٣- تعريف (الوسيط) (Median)

من أجل تعريف الوسيط لمجموعة بيانات مُعطاة سنناقش الحالتين الآتيتين:

١- إذا كانت البيانات المُعطاة خام فعندئذٍ يُعرَّف الوسيط لهذه البيانات على أنَّه القيمة التي تقع وسط البيانات بعد ترتيبها تصاعدياً (وسوف نرمز له بـ $\tilde{x}$)، فلو كانت x_1، x_2، ... و x_n بيانات خام مرتبة تصاعدياً، فإنَّ قيمة الوسيط لهذه البيانات تحسب بوساطة العلاقة الآتية:

$$\tilde{x} := \begin{cases} x_{\frac{n+1}{2}} & \text{إذا كان } n \text{ فردياً} \\ \dfrac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2} & \text{إذا كان } n \text{ زوجياً} \end{cases} \qquad [2\text{–}2]$$

٢- إذا كانت البيانات المُعطاة مجمَّعةً في جدول توزيع تكراري بـ k فئة كما في الجدول [2-1]، فعندئذٍ يُحسب الوسيط لبيانات ذلك الجدول من خلال العلاقة الآتية:

$$\tilde{x} := \tilde{L} + \frac{\dfrac{\sum_{i=1}^{n} f_i}{2} - \left(\tilde{F} - \tilde{f}\right)}{\tilde{f}} \times \mathbf{C} \qquad [2\text{–}3]$$

علماً أنَّ: $\tilde{L}$ هو الحد الأدنى للفئة (الفعليّة) الوسطيّة، وأمّا الفئة الوسطيّة فهي أول فئة تكرارها المتجمِّع الصاعد أكبر أو يساوي نصف مجموع التكرارات.

$\tilde{f}$ هو تكرار الفئة الوسطيّة.

$\tilde{F}$ هو التكرار المتجمِّع الصاعد للفئة الوسطيّة.

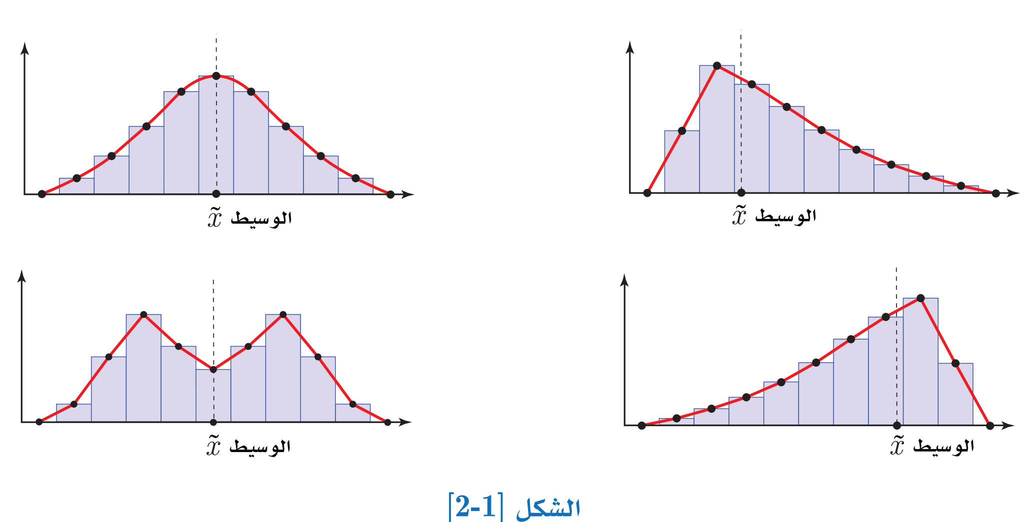
$\mathbf{C}$ هي سعة الفئة الوسطيّة (هنا لم نرفق دليل لـ $\mathbf{C}$ أو رمز خاص بها لأنَّنا نتعامل مع فئات متساوية السعة)

$\sum_{i=1}^{k} f_i$ هو مجموع التكرارات.

٢-١-٣-١- ملاحظات

١- قد يلجأ البعض إلى أخذ مركز الفئة الوسيطية كقيمة تقريبية للوسيط $\tilde{x}$ (على سبيل التبسيط)، ولكن هذا قد يؤدي إلى استنتاجات خاطئة لدى استخدامها من أجل إجراء مقارنة بينها وبين مقاييس النزعة المركزيّة الأخرى أو لدى الاستدلال بها عن شكل توزيع البيانات.

٢- كما هو واضح من تعريف الوسيط، فإنَّ الوسيط يأخذ موضعه في وسط البيانات المرتَّبة، ومن ثمَّ فإنَّ 50% من البيانات ستقع على يساره ومثل ذلك على يمينه، ومن أجل مجموعة بيانات مجمَّعة في جدول توزيع تكراري يُعبَّر عن الوسيط هندسياً على أنَّه القيمة على محور الفئات التي إذا رُسم عندها عموداً فإنَّه سيقسم التوزيع (أو المدرَّج التكراري) إلى قسمين متساويين (وليس بالضرورة منطبقين)، ويكون الانطباق في حالة التوزيعات المتناظرة فقط (للتوضيح انظر الأشكال الآتية).



الشكل [2-1]

٢-١-٣-٢- أمثلة

١- تقدَّم تسعة طلاب للاختبار النهائي في مقرَّر دراسي وكانت نتائجهم على النحو الآتي:

45	27	37	43	42	47	42	45	39

فما هي الدرجة التي تجاوزها 50% من الطلاب؟

الجواب: نلاحظ أنَّ الإجابة على هذا السؤال تتمّ بتعيين الوسيط لهذه الدرجات، ولذلك لنقم أولاً بترتيبها تصاعديا فيكون لها الترتيب الآتي:

القيم بعد ترتيبها	47	45	45	43	42	42	39	37	27
رموز القيم بعد ترتيبها	x_9	x_8	x_7	x_6	x_5	x_4	x_3	x_2	x_1

وبما أنَّ عدد البيانات فردي فإنَّه بتطبيق العلاقة [2–2] نجد أنَّ قيمة الوسيط يساوي:

$$\tilde{x} = x_{\frac{n+1}{2}} = x_5 = 42$$

إذن، فالدرجة التي تجاوزها 50% من الطلاب هي 42.

٢- بالرجوع إلى المثال (٦) من (٢-١-٢-١- الطول مقدَّراً بالسنتيمتر لخمسين طالباً-)، فنجد بعد ترتيب البيانات تصاعدياً أنَّ لها العرض الآتي:

140	153	160	162	168	171	175	179	183	189
143	155	160	165	168	172	176	182	183	190
144	157	160	166	168	173	177	182	184	191
149	159	162	167	169	175	178	182	184	197
151	160	162	168	169	175	179	182	185	199

ومنه قيمة الوسيط للبيانات الخام يساوي:

$$\tilde{x} = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2} = \frac{x_{25} + x_{26}}{2} = \frac{169 + 171}{2} = 170$$

وبعد صبّ هذه البيانات في الجدول [2-5] نجد أنَّ الفئة الرابعة هي الفئة الوسيطيّة، ومن ثمَّ تكون قيمة الوسيط لهذه البيانات يساوي:

$$\tilde{x} := \tilde{L} + \frac{\frac{\sum_{i=1}^{n} f_i}{2} - \left(\tilde{F} - \tilde{f}\right)}{\tilde{f}} \times \mathbf{C} = 160 + \frac{25 - (25 - 16)}{16} \times 10 = 170$$

لاحظ هنا أنَّ قيمة الوسيط للبيانات الخام قد توافقت مع قيمته الوسيط لبيانات جدول التوزيع التكراري الممثِّل لها.

تعقيباً على ملاحظة سابقة نلاحظ أنَّه لو أخذنا مركز الفئة الوسيطية كقيمة تقريبية لقيمة وسيط بيانات الجدول فإنَّنا سنجده يساوي 165، وهكذا نجد أنَّنا قد ارتكبنا خطأً ليس بالقليل بسبب كِبَرَ الفارق بينها وبين القيمة التي حُسِبت لوسيط بيانات جدول التوزيع.

٢-١-٣-٣- مزايا وعيوب الوسيط

١- الوسيط سهل التعريف والحساب، ولكنَّه أقل دقَّة من المتوسط.

٢- لا يتأثر بالقيم المتطرفة لأنَّه يعتمد على القيم الواقعة في الوسط.

٣- لا يعتمد على جميع القيم في حسابه، ومن ثمَّ تغيّر قيمة أو أكثر من القيم غير الواقعة في الوسط لا تؤثر في قيمة الوسيط.

٤- يمكن استخدام الوسيط إذا كانت البيانات ناقصة كأن تكون إحدى أو بعض قيم البيانات غير الواقعة في الوسط قد فُقدت لسبب ما، ولكن معلوم ترتيبها.

أخيراً نشير إلى أنَّه يُفضّل استخدام الوسيط إذا كان الاهتمام منصباً على إيجاد قيمة ممثِّلة للقيمة التي تنزع إليها البيانات بدلاً من الاهتمام بالمجموع الكلي، وكذلك عندما يكون التوزيع ملتوياً.

الآن، إذا كانت البيانات المفقودة واقعة في الوسط بعد ترتيبها تصاعدياً (ولكن معلوم موضعها) فعندئذٍ من غير الممكن استخدام المتوسط والوسيط، ولذلك يمكن اللجوء إلى المقياس الآتي لاستخدامه كمقياس للنزعة المركزيّة.

٢-١-٤- تعريف (المنوال Mode)

من أجل تعريف المنوال لمجموعة بيانات مُعطاة سنناقش الحالتين الآتيتين:

١- إذا كانت البيانات المُعطاة **خام** فعندئذٍ يُعرَّف المنوال (وسنرمز له بـ $\hat{x}$) لهذه البيانات على أنَّه القيمة الأكثر تكراراً.

٢- إذا كانت البيانات المُعطاة **مجمَّعةً** في جدول توزيع تكراري بـ k فئة كما في الجدول [2-1]، فعندئذٍ تُحسب قيمة المنوال لبيانات ذلك الجدول من خلال العلاقة الآتية:

$$\hat{x} = \hat{L} + \frac{d_1}{d_1 + d_2} C \qquad [2\text{–}4]$$

حيث لدينا: $\hat{L}$ هو الحد الأدنى للفئة المنوالية،

d_1 هو الفرق بين تكرار الفئة المنواليّة وتكرار الفئة السابقة لها مباشرةً،

d_2 هو الفرق بين تكرار الفئة المنواليّة وتكرار الفئة اللاحقة بها مباشرةً،

C هو سعة الفئة المنواليّة.

علماً أنَّ الفئة المنواليّة هي تلك الفئة التي تكرارها أكبر من تكرار الفئة السابقة لها واللاحقة بها مباشرةً، وعلى ألا تكون هـذه الفئـة طرفيّة. أي أنَّ الفئة الأولى والأخيرة في التوزيع التكراري لا ينظر إليهما كفئات منواليّه.

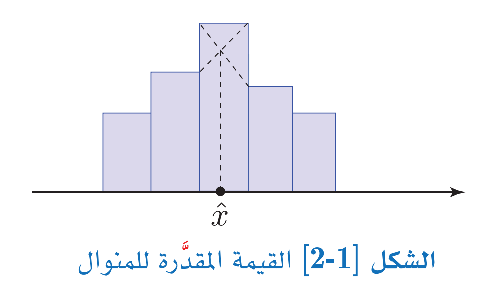
٢-١-٤-١- ملاحظات

١- قد يلجأ البعض إلى أخذ **مركز** الفئة المنواليّة كقيمة تقريبية للمنوال $\hat{x}$، ولكن هذا قد يؤدي إلى استنتاجات خاطئة لدى استخدامها من أجل إجراء مقارنة بينها وبين مقاييس النزعة المركزيّة الأخرى أو لدى الاستدلال بها عن شكل توزيع البيانات.

٢- من تعريف المنوال يُلاحظ أنَّ المنوال قد لا يكون موجوداً في حال أنَّ جميع البيانات مختلفة عن بعضها البعض الآخر أو أنَّ لجميع القيم الممثِّلة التكرار نفسه. كذلك يُلاحظ أنَّه حتى في حال وجوده قد لا يكون وحيداً في حال تساوي التكرار لقياسين على الأقل من هذه القياسات حيث يكون لدينا منوالين على الأقل في هذه الحالة.

٣- يمكن تعيين المنوال هندسياً إذا كان المدرَّج التكراري مُعطى، وذلك على النحو الآتي:

- نصل بقطعة مستقيمة بين النهاية اليمنى لقمة الفئة الفعليَّة المنوالية مع النهاية اليمنى لقمَّة الفئة السابقة لها.
- نصل بقطعة مستقيمة بين النهاية اليسرى لقمة الفئة الفعليَّة المنوالية مع النهاية اليسرى لقمَّة الفئة اللاحقة بها.
- نسقط عمود من نقطة تقاطع المستقيمين الناشئين على محور الفئات فتكون النقطة الموافقة لهذا المسقط هي قيمة للمنوال.



الشكل [2-1] القيمة المقدَّرة للمنوال

إنَّ هذه الطريقة لتعيين المنوال تحتاج إلى دقة في الرسم واستخدام مقاييس دقيقة، ولذلك لا تستخدم إلاَّ عند الضرورة.

◀ ٢-١-٤-٢- أمثلة

١- لدى وزن عشرة صناديق متماثلة تحتوي على برتقال وجدنا أوزانها كما يلي:

8	8.5	8.25	7.85	8	7.75	7.95	8	8.45	8.25

فما هو منوال هذه البيانات؟

✍ **الجواب:** نلاحظ أنَّ القيمة الأكثر تكراراً بين البيانات المُعطاة هي 8 ، وهي قيمة وحيدة ومن ثمَّ لدينا منوال وحيد في هذه البيانات هو $\hat{x} = 8$.

٢- في سباق للسيَّارات عالية السرعة في حلبة محدَّدة (**وليكن على سبيل المثال الفورمولا**) دوِّنت الأزمنة الآتية لثمانية متسابقين (**الزمن مقدَّر بالدقيقة**):

13.62	12.79	12.88	13.09	13.00	12.95	13.82	13.57

فما هو منوال هذه البيانات؟

✍ **الجواب:** نلاحظ هنا أنَّ لجميع القيم التكرار نفسه، ومن ثمَّ لا يوجد منوال لهذه البيانات.

٣- بالرجوع إلى المثال (٦) من (٢-١-٢-١- **الطول مقدَّراً بالسنتيمتر لخمسين طالباً**-)، فنجد أنَّ في تلك البيانات أربع قيم لها أعلى تكرار وكل واحدة منها تكرَّرت أربع مرَّات، ولذلك لدينا في تلك البيانات أربعة مناويل هي $\hat{x}_1 = 160$ و $\hat{x}_2 = 168$ و $\hat{x}_3 = 175$ و $\hat{x}_4 = 182$.

وبعد صبّ تلك البيانات في الجدول [2-4] نجد أنَّ الفئة الثالثة هي الفئة المنوالية، ومن ثمَّ تكون قيمة المنوال لتلك البيانات يساوي:

$$\hat{x} = \hat{L} + \frac{d_1}{d_1 + d_2} C = 160 + \frac{11}{11+5} \times 10 = 166.88$$

لاحظ هنا أنَّه لو أخذنا مركز الفئة المنوالية كقيمة تقريبيّة لقيمة منوال بيانات الجدول فإنَّنا سنجده يساوي 165، وهذه القيم تقلّ بمقدار 1.88 عن القيمة المحسوبة آنفاً، وكذلك لا تتوافق مع أية قيمة من قيم المناويل للبيانات الخام نفسها. إذن فمن غير المقبول استخدام مركز الفئة المنوالية كقيمة تقريبية لقيمة منوال بيانات جدول التوزيع التكراري.

٢-١-٤-٢- مزايا وعيوب المنوال

١- إنَّ المنوال هو أقل مقاييس النزعة المركزيّة استخداماً، ويعدُّ غير ذي جدوى عمليّاً إذا كان عدد البيانات قليلاً (هذا إنَّ وجِدَ أصلاً)، وأمَّا في حالة البيانات كبيرة العدد ومن أجل جداول التوزيع التكرارية فيمكن الاستفادة منه لكون تأثرِه بتغيّر بعض قيم البيانات قد لا يكون كبيراً بالقدر الذي يُلغي استخدامه.

٢- من فوائد المنوال أنَّه يمكن استخدامه في حالة البيانات النوعيّة (غير العددية)، فعلى سبيل المثال إذا كان لون العيون لمعظم النَّاس في بلد ما هو اللون البني، فعندئذٍ يمكننا القول إنَّ اللون المنوالي لعيون هؤلاء النَّاس هو اللون البني.

٣- من العيوب الكبيرة للمنوال أنَّه قد لا يكون وحيداً، وهذا يعني أنَّه لا يمكن تعريفه بشكل وحيد كمقياس للنزعة المركزيّة.

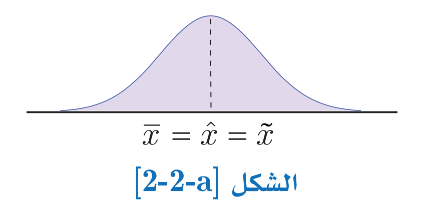
٤- بالرغم من أنَّ استخدام المنوال قليلٌ جداً، ونادراً ما يُلجأ إليه أيضاً، ولكن في حال فقدان بعض البيانات الواقعة في الوسط (بعد ترتيبها تصاعدياً أو تنازلياً) فإنَّ للمنوال دور جيد في تعيين القيمة الممثِّلة لتمركز البيانات. كما أنَّه يشير في كثير من الحالات إلى عدد العوامل المؤثِّرة في توليد البيانات من خلال عدد المناويل التي ستظهر للبيانات عند تعيينها.

٢-١-٥- ملاحظة ختامية

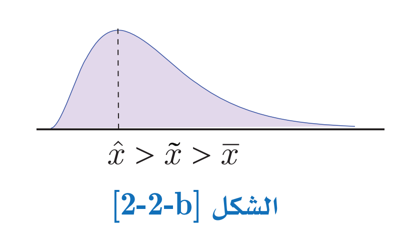
من أجل التوزيعات التكرارية المتماثلة يكون لدينا $\bar{x} = \hat{x} = \tilde{x}$، وأمَّا إذا كان التوزيع قريب من التماثل فإنَّه يكون $\bar{x} \approx \hat{x} \approx \tilde{x}$، وفي هذه الحالة وجد تجريبياً أنَّ:

$$\bar{x} - \hat{x} = 3(\bar{x} - \tilde{x})$$

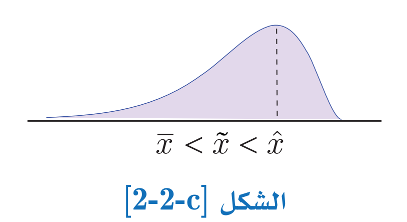
وأمّا التوضُّع النسبي لقيم المتوسط والوسيط والمنوال في حالة التوزيعات المختلفة تُظهرها لنا الأشكال الآتية:



الشكل [2-2-a]



الشكل [2-2-b]



الشكل [2-2-c]

٢ - ٢
الربيعيَّات

لقد قدَّمنا فيما سبق بعض مقاييس النزعة المركزيّة، حيث لاحظنا أنَّ لكل مقياس من هذه المقاييس موضع يشغله على محور القيم أو الفئات، ومن ثمَّ هذه المقاييس التي قدِّمت هي في الواقع مقاييس موضع أيضاً. لكن يوجد في الواقع مقاييس موضع أخرى لا تنتمي لمجموعة مقاييس النَّزعة المركزيّة، ومنها على سبيل المثال الربيعيَّات والمئينات التي سنقدِّمها تباعاً، ولكن من أجل البيانات الكميّة الخام فقط وبشيء من الإيجاز أيضاً.

٢-٢-١- تعريف (الربيعيَّات)

لتكن x_1 و x_2 و... و x_n بيانات **مرتبة تصاعدياً**. عندئذٍ تُعرَّف الربيعيَّات على أنَّها تلك القيم التي تقسم البيانات المُعطاة إلى أربع شرائح (**أو أجزاء**) متساوية الأعداد من البيانات (**أي في كل شريحة العدد نفسه من قيم البيانات**).

٢-٢-١-١ ملاحظات

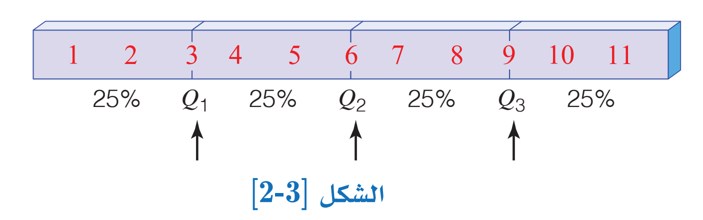
من هذا التعريف يتضّح لنا وجود ثلاثة ربيعيَّات وهي:

١- القيمة التي يقع قبلها 25% وبعدها 75% من البيانات المرتبة وتُدعى **الرُبيْعي الأول** First Quartile (**أو الرُبيْعي الأدنى** Lower Quartile)، ويُرمز له بالرمز $\mathbf{Q}_1$.

٢- القيمة التي يقع قبلها 50% وبعدها 50% من البيانات المرتبة وتُدعى **الرُبيْعي الثاني** Second Quartile، ويُرمز له بـ $\mathbf{Q}_2$، ويُلاحظ هنا أنَّ الرُبيْعي الثاني هو الوسيط نفسه، أي أنَّ $\mathbf{Q}_2 = \tilde{x}$.

٣- القيمة التي يقع قبلها 75% وبعدها 25% من البيانات المرتبة وتُدعى **الرُبيْعي الثالث** Third Quartile (**أو الرُبيْعي الأعلى** Upper Quartile)، ويُرمز له بـ $\mathbf{Q}_3$.

ومن أجل مجموعة البيانات $A = \{1,2, \cdots ,11\}$ يمكننا تقديم العرض الهيكلي الآتي للرُبيْعيات الثلاث.



الشكل [2-3]

٢-٢-٢- تعيين الربيعيَّات

في الواقع توجد طرائق عديدة لتعيين الربيعيَّات، وهذه الطرائق تتفاوت في دقّة نتائجها وفقاً للقاعدة الرياضياتية المستخدمة في الحساب، وهذا بدوره يجعل نتائج حساب الربيعيَّات باستخدام البرامج الإحصائية تختلف من برنامج إلى آخر. في كتابنا هذا سنقدّم إحدى هذه الطرائق التي تعدّ من أكثر الطرائق دقةً، وهي مستخدمة في الكثير من البرامج الإحصائية أيضاً.

لتكن x_1 و x_2 و... و x_n بيانات خام **مرتبة تصاعدياً**، ولنضع بالتعريف:

$$q_r := \frac{r\,(n+1)}{4} \quad ; r = 1, 2, 3 \qquad [2\text{–}5]$$

والذي يُدعى **رتبة** **Rank** الرُبيْعي r (**وهذه القيمة تحدِّد موضع الرُبيْعي بين البيانات المرتّبة**)، فعندئذٍ تُعيَّن قيمة الرُبيْعي $\mathbf{Q}_r$ مع $r = 1, 2, 3$ على النحو الآتي:

بفرض أنَّ k هو الجزء الصحيح من العدد q_r، وأنَّ الباقي من هذا العدد يساوي s، فعندئذٍ يُحسب الرُبيْعي $\mathbf{Q}_r$ من خلال العلاقة الآتية:

$$\mathbf{Q}_r = x_k + s\,(x_{k+1} - x_k) \quad ; r = 1, 2, 3 \qquad [2\text{–}6]$$

لاحظ هنا أنّه إذا كان الباقي $s = 0$ أو كان $x_{k+1} = x_k$ فعندئذٍ سيكون لدينا $\mathbf{Q}_r = x_k$.

◄ ٢-٢-٢-١- أمثلة

١- لدينا درجات اختبار نهائي (الدرجة القصوى 50) لخمس وعشرين طالباً كما هو آتٍ.

48	37	25	45	44
44	49	29	42	48
45	50	39	36	47
41	48	38	15	47
43	47	39	25	48

ولنقم بتعيين الربيعيَّات الثلاثة $\mathbf{Q}_1$ و $\mathbf{Q}_2$ و $\mathbf{Q}_3$.

الأجوبة: من أجل ذلك لنقم أولاً بترتب البيانات التي لدينا تصاعدياً، فنجد لها التسلسل الآتي:

15	37	42	45	48
25	38	43	47	48
25	39	44	47	48
29	39	44	47	49
36	41	45	48	50

وبما أنَّ $n = 25$ فإنَّ رتبة الرُّبيْعي Q_1 ، Q_2 و Q_3 ستكون على الترتيب هي:

$$q_1 = \frac{n+1}{4} = \frac{25+1}{4} = 6.5$$

$$q_2 = \frac{2(n+1)}{4} = \frac{2(25+1)}{4} = 13$$

$$q_3 = \frac{3(n+1)}{4} = \frac{3(25+1)}{4} = 19.5$$

فنجد من أجل الرُبيْعي الأول Q_1 (أي أنَّ $r = 1$) أنَّ الجزء الصحيح من العدد q_1 هو $k = 6$، وأمّا الباقي من هذا العدد يساوي $S = 0.5$، ومن ثمَّ يكون لدينا:

$$Q_1 = x_k + s(x_{k+1} - x_k) = x_6 + 0.5(x_7 - x_6)$$
$$= 37 + 0.5\ (38 - 37) = 37.5$$

أي أنَّ 25% من الطلاب قد حصلوا على درجة دون 37.5، وأمَّا الباقي (وهم 75%) فقد حصلوا على درجة أعلى من 37.5 في هذا الاختبار.

ومن أجل الرُبيْعي الثاني Q_2 نجد أنَّ الجزء الصحيح من العدد q_2 هو $k = 13$، وأمّا الباقي من هذا العدد يساوي $S = 0$، ومن ثمَّ يكون لدينا:

$$Q_2 = x_k + s(x_{k+1} - x_k) = x_{13} = 44$$

أي أنَّ 50% من الطلاب قد حصلوا على درجة دون 44، وأمَّا الباقي (وهم 50% أيضاً) فقد حصلوا على درجة أعلى من 44 في هذا الاختبار.

وأخيراً من أجل الرُبيْعي الثالث Q_3 نجد أنَّ الجزء الصحيح من العدد q_3 هو $k = 19$، وأمّا الباقي من هذا العدد يساوي $S = 0.5$، ومن ثمَّ يكون لدينا:

$$Q_3 = x_k + s(x_{k+1} - x_k) = x_{19} + 0.5(x_{20} - x_{19})$$
$$= 47 + 0.5\ (48 - 47) = 47.5$$

أي أنَّ 75% من الطلاب قد حصلوا على درجة دون 47.5، وأمَّا الباقي (وهم 25%) فقد حصلوا على درجة أعلى من 47.5 في هذا الاختبار.

٢- لدى معاينة مجموعة مكوَّنة من 20 رجلاً بالغاً من أجل معرفة أوزانهم وجدنا البيانات الآتية (مقدَّرةً بالكيلو غرام):

86	95	66	76
87	98	77	92
85	98	78	65
73	106	77	88
70	115	77	102

ولنقم بتعيين الربيعيَّات الثلاثة Q_1 و Q_2 و Q_3.

الأجوبة: من أجل ذلك لنقم أولاً بترتب البيانات التي لدينا تصاعدياً، فنجد لها التسلسل الآتي:

65	77	86	98
66	77	87	98
70	77	88	102
73	78	92	106
76	85	95	115

وبما أنَّ $n = 20$ فإنَّ رتبة الرُبيْعي Q_1 ، Q_2 و Q_3 ستكون على الترتيب هي:

$$q_1 = \frac{n+1}{4} = \frac{20+1}{4} = 5.25$$

$$q_2 = \frac{2(n+1)}{4} = \frac{2(20+1)}{4} = 10.5$$

$$q_3 = \frac{3(n+1)}{4} = \frac{3(20+1)}{4} = 15.75$$

فنجد من أجل الرُبيْعي الأول Q_1 (أي أنَّ $r = 1$) أنَّ الجزء الصحيح من العدد q_1 هو $k = 5$، وأمّا الباقي من هذا العدد يساوي $S = 0.25$، ومن ثمَّ يكون لدينا:

$$\begin{aligned} Q_1 &= x_k + s(x_{k+1} - x_k) = x_5 + 0.25(x_6 - x_5) \\ &= 76 + 0.25\,(77 - 76) = 76.25 \end{aligned}$$

أي أنَّ 25% من هؤلاء الرجال لهم أوزان دون 76.25، وأمَّا الباقي (وهم 75%) فإنَّ أوزانهم أكثر من 76.25.

ومن أجل الرُبيْعي الثاني Q_2 نجد أنَّ الجزء الصحيح من العدد q_2 هو $k = 10$، وأمّا الباقي من هذا العدد يساوي $S = 0.5$، ومن ثمَّ يكون لدينا:

$$\begin{aligned} Q_2 &= x_k + s(x_{k+1} - x_k) = x_{10} + 0.5(x_{11} - x_{10}) \\ &= 85 + 0.5\,(86 - 85) = 85.5 \end{aligned}$$

أي أنَّ 50% من هؤلاء الرجال لهم أوزان دون 85.5، وأمَّا النصف الباقي منهم فلهم أوزان أكثر من 85.5.

وأخيراً من أجل الرُبيْعي الثالث Q_3 نجد أنَّ الجزء الصحيح من العدد q_3 هو $k = 15$، وأمّا الباقي من هذا العدد يساوي $S = 0.75$، ومن ثمَّ يكون لدينا:

$$\begin{aligned} Q_3 &= x_k + s(x_{k+1} - x_k) = x_{15} + 0.5(x_{16} - x_{15}) \\ &= 95 + 0.75\,(98 - 95) = 97.25 \end{aligned}$$

أي أنَّ 75% من هؤلاء الرجال لهم أوزان دون 97.25، وأمَّا الباقي (وهم 25%) فإنَّ أوزانهم أكثر من 97.25.

٢-٢-٢-٢- ملاحظات

١- تجدر الإشارة هنا إلى أنَّ قيمة الرُّبيْعي قد لا تكون موجودةً بين قيم البيانات المُعطاة.

٢- إنَّ القيمتين Q_1 و Q_3 ليستا من مقاييس النَّزعة المركزيّة.

٣- من الاستخدامات المهمّة للرُّبيْعيات تحديد ما إذا كانت قيمة x من مجموعة بيانات مُعطاة هي قيمة متطرفة Extreme Value (أو قيمة منعزلة Outlier Value) أم لا. حيث يُقال عن قيمة x من مجموعة بيانات مُعطاة إنَّها متطرفةٌ إذا حقَّقت إحدى العلاقتين الآتيتين:

$$x < Q_1 - 1.5(Q_3 - Q_1) \qquad [2\text{–}7\text{-}a]$$

$$x > Q_3 + 1.5(Q_3 - Q_1) \qquad [2\text{–}7\text{-}b]$$

فإذا كانت العلاقة [2–7-a] محقَّقةً، فعندئذٍ يُقال إنَّ القيمة x متطرفة بصغرها، وأمَّا المقدار $Q_1 - 1.5(Q_3 - Q_1)$ فإنَّه يُدعى أدنى حاجز Lowest Fence (لأنَّه حاجز يفصل بين القيم المتطرفة بصغرها وباقي القيم غير المتطرفة)، وسنرمز لها بـ **LF**، أي أنَّه لدينا:

$$LF = Q_1 - 1.5(Q_3 - Q_1)$$

وهو يمثِّل في الواقع **الحد الأعلى للقيم المتطرفة بصغرها** من أجل مجموعة بيانات مُعطاة.

وأمّا إذا كانت العلاقة [2–7-b] محقَّقةً، فعندئذٍ يُقال إنَّ القيمة x متطرفة بكبرها، وحينئذ يُقال عن المقدار $Q_3 + 1.5(Q_3 - Q_1)$ إنَّه أعلى حاجز Highest Fence (لأنَّه حاجز يفصل بين القيم المتطرفة بكبرها وباقي القيم غير المتطرفة)، وسنرمز له بـ **HF**، أي أنَّه لدينا:

$$HF = Q_3 + 1.5(Q_3 - Q_1)$$

وهو يمثِّل في الواقع **الحد الأدنى للقيم المتطرفة بكبرها** من مجموعة بيانات مُعطاة.

فعلى سبيل المثال لو أخذنا البيانات الآتية:

قيمة متطرّفة ← 15

6	7	15	4	7	5	5	4	5	3	4	5
7	3	4	1	4	5	4	4	7	5	5	6
5	7	1	6	4	7						

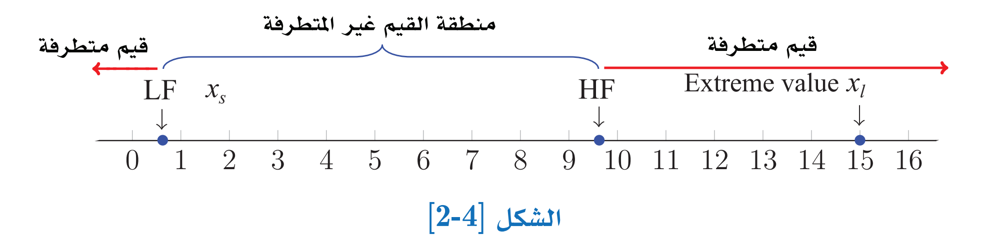
فنجد أنَّ $Q_1 = 4$ و $Q_3 = 6.25$، وكذلك:

$$Q_3 - Q_1 = 6.25 - 4 = 2.25$$

ومن ثمَّ يكون لدينا:

$$LF = Q_1 - 1.5(Q_3 - Q_1) = 4 - (1.5 \times 2.25) = 0.625 < 1$$

$$HF = Q_3 + 1.5(Q_3 - Q_1) = 6.25 + (1.5 \times 2.25) = 9.625 < 15$$



الشكل [2-4]

فنجد أنَّ القيمة 15 **متطرفة** بكبرها ولا وجود لقيم متطرفة بصغرها.

٢ - ٣
المئينات

يعدّ المئين من مقاييس الموضع المهمّة وذلك لأنَّه يوفر معلومات حول كيفية امتداد البيانات من أصغر قيمة إلى أكبر قيمة، فعلى سبيل المثال يُذكر في كثير من الأحيان في جدول نتائج تقديم الطلبات لمؤسسة ما أنَّ 85% من المتقدمين استوفوا الشروط المطلوبة للتعيين فقط. أو أن يُقال إنَّ 72% فقط من الطلاب حققّوا درجة النجاح في المقرّر X.

٢-٣-١ تعريف (المئينات):

لتكن x_1 و x_2 و... و x_n بيانات **مرتبة تصاعدياً**. عندئذٍ تُعرَّف المئينات على أنَّها تلك القيم التي تقسم البيانات إلى مئة شريحة (**أو جزء**) متساوية الأعداد من البيانات (**أي في كل شريحة العدد نفسه من قيم البيانات**).

٢-٣-٢- ملاحظات

١- من هذا التعريف يتضّح لنا وجود تسعة وتسعين **مئيناً**، علماً أنَّه من أجل $r = 1, 2, \cdots, 99$ فإنَّ القيمة التي يقع قبلها $r\%$ من البيانات (**المرتبة تصاعدياً**) وبعدها $(100 - r)\ \%$ تُدعى **المئين** ذي الرقم r (rth Percentile أو **المئين الرائد** r-Percentile)، ويُرمز لها بـ $\mathbf{P}_r$.

٢- من تعريف المئين نلاحظ أنَّ $\mathbf{P}_{25} = \mathbf{Q}_1$ و $\mathbf{P}_{50} = \mathbf{Q}_2 = \tilde{x}$ و $\mathbf{P}_{75} = \mathbf{Q}_3$.

٣- بما أنَّ عدد المئينات كبير (**لدينا 99 مئيناً**) فإنَّ هذا النوع من المقاييس يكون أكثر وضوحاً كلما ازداد عدد البيانات، ولذلك تكون نتائج هذا النوع من المقاييس أكثر دقّة عندما يكون عدد البيانات كبيراً، وما سنقدّمه من أمثلة على بيانات قليلة العدد سيكون من باب التوضيح والتبسيط.

٢-٣-٣- تعيين المئينات

إنَّ طريقة تعيين المئينات تُماثل طريقة تعيين الربيعيَّات تماماً، فلو كانت x_1 و x_2 و... و x_n بيانات **مرتبة تصاعدياً**، فإنَّنا نضع بالتعريف:

$$p_r = \frac{r\ (n+1)}{100} \quad ; r = 1, 2, \ldots, 99 \qquad [2\text{–}8\text{-a}]$$

وهذه القيمة p_r تُدعى **رتبة** المئين r في البيانات المرتبة تصاعدياً، وعملها تحديد موضع هـذا المئـين بـين البيانات المرتّبة تصاعدياً. عندئذٍ لتعيين قيمة المئين $\mathbf{P}_r$ نقوم بما يلي:

بفرض أنَّ k هو الجزء الصحيح من المقدار p_r، وأنَّ الباقي منه يساوي s، فعندئذٍ تُحسب قيمة المئين $\mathbf{P}_r$ من خلال العلاقة الآتية:

$$\mathbf{P}_r = x_k + s(x_{k+1} - x_k) \quad ; r = 1, 2, ..., 99 \qquad [2\text{–}8\text{-b}]$$

نلاحظ هنا أنَّه إذا كان $s = 0$ أو $x_{k+1} = x_k$ فإنَّه سيكون لدينا $\mathbf{P}_r = x_k$ فقط.

◀ ٢-٣-٢-١- مثال

تقدَّم لنا البيانات الآتية الطول لخمسين طالباً في عمادة السّنة الأولى المشتركة.

185	169	173	179	173	178	147	171	173	172
169	201	184	160	163	189	190	158	149	173
172	149	177	158	169	172	171	168	181	178
195	181	171	195	190	167	185	173	180	179
169	181	172	196	152	175	179	169	175	175

ولنقم بتعيين المئين $\mathbf{P}_{15}$ ، $\mathbf{P}_{43}$ و $\mathbf{P}_{87}$ لهذه البيانات.

الأجوبة: من أجل ذلك لنقم بترتيب البيانات المُعطاة تصاعدياً، فنجد لها التسلسل الآتي:

147	158	169	171	172	173	177	179	184	190
149	160	169	171	172	173	178	180	185	195
149	163	169	171	173	175	178	181	185	195
152	167	169	172	173	175	179	181	189	196
158	168	169	172	173	175	179	181	190	201

البيانات بعد ترتيبها تصاعدياً

وبما أنَّ $n = 50$ فإنَّ رتبة المئين $\mathbf{P}_{15}$ ، $\mathbf{P}_{43}$ و $\mathbf{P}_{87}$ ستكون على الترتيب هي:

$$p_{15} = \frac{15(n+1)}{100} = \frac{15(50+1)}{100} = 7.65$$

$$p_{43} = \frac{43(n+1)}{100} = \frac{43(50+1)}{100} = 21.93$$

$$p_{87} = \frac{87(n+1)}{100} = \frac{87(50+1)}{100} = 44.37$$

ومن ثمَّ يكون لدينا:

$$\mathbf{P}_{15} = x_k + s(x_{k+1} - x_k) = x_7 + 0.65(x_8 - x_7)$$
$$= 160 + 0.65(163 - 160) = 161.95$$

$$\mathbf{P}_{43} = x_k + s(x_{k+1} - x_k) = x_{21} + 0.93(x_{22} - x_{21})$$
$$= 172 + 0.93(172 - 172) = 172$$

$$\mathbf{P}_{87} = x_k + s(x_{k+1} - x_k) = x_{44} + 0.37(x_{45} - x_{44})$$
$$= 189 + 0.37(190 - 189) = 189.37$$

٢- لنقم بحساب المئينات $\mathbf{P}_{25}$ ، $\mathbf{P}_{50}$ و $\mathbf{P}_{75}$ لبيانات المثال (٢) من (٢-٢-٢-١)، حيث قمنا بحسابها بوساطة الربيعيَّات سابقاً، علماً أنَّ البيانات المرتَّبة تصاعدياً كانت كما يأتي:

65	77	86	98
66	77	87	98
70	77	88	102
73	78	92	106
76	85	95	115

الأجوبة: بما أنَّ $n = 20$ فإنَّ رتبة المئين $\mathbf{P}_{25}$ ، $\mathbf{P}_{50}$ و $\mathbf{P}_{75}$ ستكون على الترتيب هي:

$$p_{25} = \frac{25(n+1)}{100} = \frac{25(20+1)}{100} = 5.25$$

$$p_{50} = \frac{50(n+1)}{100} = \frac{50(20+1)}{100} = 10.5$$

$$p_{75} = \frac{75(n+1)}{100} = \frac{75(20+1)}{100} = 15.75$$

ومن ثمَّ يكون لدينا:

$$\mathbf{P}_{25} = x_5 + 0.25\,(x_6 - x_5) = 76 + 0.25(77 - 76) = 76.25$$

$$\mathbf{P}_{50} = x_{10} + 0.5\,(x_{11} - x_{10}) = 85 + 0.5(86 - 85) = 85.5$$

$$\mathbf{P}_{75} = x_{15} + 0.75(x_{16} - x_{15}) = 95 + 0.75(98 - 95) = 97.25$$

وبالرجوع إلى قيم الربيعيَّات الثلاثة التي قمنا بحسابها في ذلك المثال $\mathbf{Q}_1 = 76.25$ ، $\mathbf{Q}_2 = 85.5$ و $\mathbf{Q}_3 = 97.25$ نجد تطابق نتائجها مع قيم المئينات $\mathbf{P}_{25}$ ، $\mathbf{P}_{50}$ و $\mathbf{P}_{75}$ بشكل تام.

٢-٣-٢-٢- ملاحظات

١- توجد طريقة تقريبية ولكنها مقبولة لتعيين المئين الموافق لأيّة قيمة من قيم بيانات خام مُعطاة، وذلك على النحو الآتي.

بفرض أنَّ x_1 و x_2 و... و x_n بيانات **مرتبة تصاعدياً**، ومن أجل $1 \le i \le n$ لنرمز بـ $\mathbf{P}_{x_i}$ لقيمة المئين الموافقة للقيمة x_i، وبـ N_i لعدد البيانات التي هي أصغر من x_i، فعندئذٍ تُعطى قيمة $\mathbf{P}_{x_i}$ بالعلاقة الآتية:

$$\mathbf{P}_{x_i} := \frac{N_i + 0.5}{n} \times 100 \qquad [2\text{-}9]$$

فعلى سبيل المثال لو أخذنا البيانات الآتية:

2, 9, 2, 6, 4, 5, 3, 8, -1, 0, 6

وبعد ترتيب هذه البيانات تصاعدياً نجد المئين الموافق لكل قيمة من القيم المرتبة كما في الجدول الآتي (قمنا باستخدام الجدول على سبيل التوضيح فقط).

الجدول [2-4]

x_i	-1 x_1	0 x_2	2 x_3	2 x_4	3 x_5	4 x_6	5 x_7	6 x_8	6 x_9	8 x_{10}	9 x_{11}
N_i	0	1	2	3	4	5	6	7	8	9	10
$\mathbf{P}_{x_i}$	4.5	13.6	22.7	31.8	40.9	50	59.1	68.2	77.3	86.4	95.5

فلو أردنا على سبيل المثال تطبيق الطريقة السابقة لتعيين المئين $\mathbf{P}_{40}$ للبيانات المعطاة فإنَّنا سنجد رتبة المئين $\mathbf{P}_{40}$ تساوي:

$$p_{40} = \frac{40(n+1)}{100} = \frac{40(11+1)}{100} = 4.80$$

ومن ثمَّ تكون قيمة المئين $\mathbf{P}_{40}$ هي:

$$\mathbf{P}_{40} = x_4 + 0.80\,(x_5 - x_4) = 2 + 0.80\,(3-2) = 2.80$$

فنجد أنَّها تتوافق مع نتيجة الطريقة الأخيرة بتقريبٍ جيد حيث كان لدينا $x_5 = 3$ والمئين الموافق لها يساوي $\mathbf{P}_{x_5} = 40.9$، فلو نظرنا إلى النتيجة السابقة فإنَّنا سنلاحظ أنَّها تشير إلى المئين $\mathbf{P}_{40}$ في البيانات، وموضعها على محور القيم يأتي قُبَيْل النقطة الممثَّلة بالقيمة $x_5 = 3$ (والتي هي 2.80).

٢ - ٤
الأعداد الخمسة والتمثيل الصندوقي للبيانات

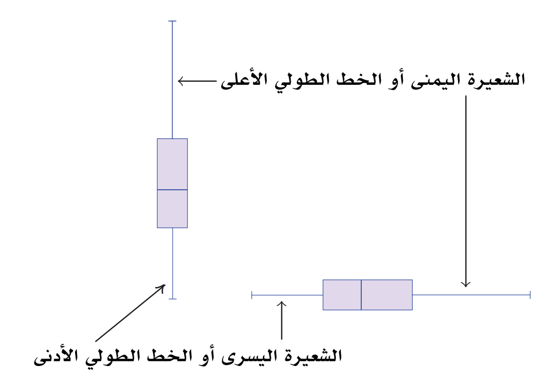
يقصد بالأعداد الخمسة في الإحصاء الوصفي القيم الآتية:

١- أصغر قيمة في البيانات x_s،

٢- أكبر قيمة في البيانات x_ℓ،

٣- قيم الربيعيَّات الثلاثة $\mathbf{Q}_1$ و $\mathbf{Q}_2$ و $\mathbf{Q}_3$.

إنَّ هذه الأعداد الخمسة تساعدنا في وصف تمركز البيانات وانتشارها وشكل توزيعها، وأمّا العرض (أو التمثيل) الصندوقي Box Plot للبيانات فهو عرض رسومي للبيانات يقوم على أساس مُلخَّص الأعداد الخمسة السابقة، ويتكوَّن من صندوق وخطين طوليين يتوسطانه على طرفيه يُمنةً ويُسرى (أو من الأعلى والأدنى) يُدعيان "الشُعَيْرتان Whiskers". ويُقدَّم هذا العرض وفقاً لأحد الشكلين الآتيين:



الشكل [2-14]

أمَّا نهاية الشُعَيْرة اليمنى (وسنرمز لها بـ x_e) واليسرى (وسنرمز لها بـ x_a) فإنَّهما يعيّنان كما يلي:

١- إذا كانت مجموعة البيانات لا تحتوي قيم متطرفة، فعندئذٍ يكون $x_\mathrm{a} = x_s$ و $x_\mathrm{e} = x_\ell$.

٢- إذا كانت مجموعة البيانات تحتوي على قيم متطرفة بصغرها فقط، فعندئذٍ يكون لدينا $\mathbf{LF} = x_\mathrm{a}$. أي أنَّ نهاية الشُعيرة اليسرى x_a توافق الحد الأعلى للقيم المتطرفة بصغرها، ومن ثمَّ تبلغ الشُعيرة اليسرى الطول الأعظمي لها. أمَّا x_e فيكون مساوياً لـــ x_ℓ في هذه الحالة.

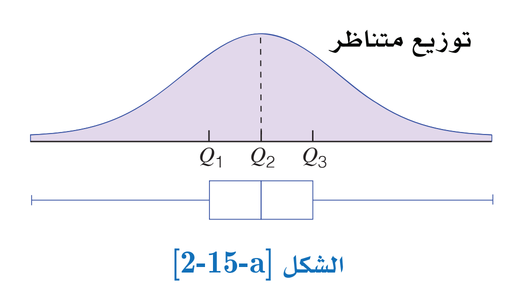
٣- إذا كانت مجموعة البيانات تحتوي على قيم متطرفة بكبرها فقط، فعندئذٍ يكون لدينا $\mathbf{HF} = x_\mathrm{e}$. أي أنَّ نهاية الشُعيرة اليمنى x_e توافق الحد الأدنى للقيم المتطرفة بكبرها، ومن ثمَّ تبلغ الشُعيرة اليمنى الطول الأعظمي لها. أمَّا x_a فيكون مساوياً لـــ x_s في هذه الحالة.

٤- إذا كانت مجموعة البيانات تحتوي على قيم متطرفة بصغرها وأخرى متطرفة بكبرها، فعندئذٍ يكون لدينا $x_a = \mathbf{LF}$ و $x_e = \mathbf{HF}$، وفي هذه الحالة تبلغ الشُعيرتان اليمنى واليسرى الطول الأعظمي لهما.

٢-٤-١- ملاحظات

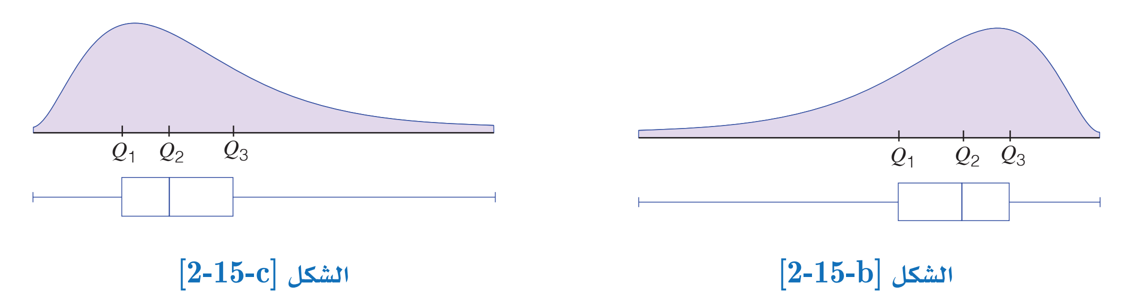
١- عند تقديم الرسم الصندوقي يُشار إلى كل قيمة متطرفة بنجمة (*) أو نقطة (•).

٢- إذا كان توزيع البيانات متناظراً، فعندئذٍ سيتمركز الصندوق وخط الوسيط بين نقطتي النهاية للشُعيرّتين (انظر الشكل التوضيحي الآتي).



الشكل [2-15-a]

٣- إذا كان توزيع البيانات ملتوياً نحو اليسار فعندئذٍ سينزاح الصندوق وخط الوسيط لليمين باتجاه نهاية الشُعيرة اليمنى، وأمّا إذا كان توزيع البيانات ملتوياً نحو اليمين فإنَّ الصندوق وخط الوسيط سينزاحان إلى اليسار باتجاه نهاية الشُعيرة اليسرى (انظر الشكلين التوضيحيين الآتيين).



الشكل [2-15-b]　　　الشكل [2-15-c]

◀ ٢-٦-١- مثال

١- لتكن لدينا مجموعة البيانات الآتية:

$$8, 4, 7, 6, 9, 1, 5, 9, 3, 17, 5, -7$$

فنجد بعد ترتيبها تصاعدياً أنّ لها العرض الآتي:

$$-7, 1, 3, 4, 5, 5, 6, 7, 8, 9, 9, 17$$

ومن ثمَّ تكون الأعداد الخمسة لهذه البيانات هي:

$$x_s = -7 \quad \& \quad x_\ell = 17 \quad \& \quad \mathbf{Q}_1 = 3.25 \quad \& \quad \tilde{x} = \mathbf{Q}_2 = 5.5 \quad \& \quad \mathbf{Q}_3 = 8.75$$

وأمَّا من أجل تحديد القيم المتطرفة بصغرها والتي تكون أصغر من نهاية الشُعيرة اليسرى، فلدينا:

$$\mathbf{LF} = 3.25 - 1.5(8.75 - 3.25) = -5$$

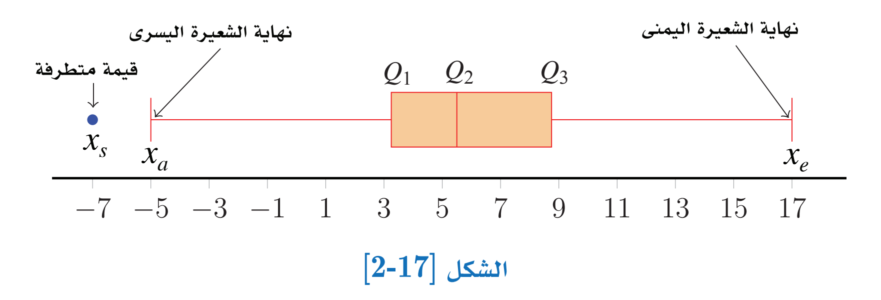
أي أنَّ القيم المتطرفة بصغرها هي تلك القيم التي تكون أصغر من 5- ، ومن ثمَّ لدينا قيمة متطرفة بصغرها في البيانات المقدَّمة هي $x_s = -7$.

كذلك من أجل تحديد القيم المتطرفة بكبرها التي تكون أكبر من نهاية الشُعيرة اليمنى لدينا:

$$\mathbf{HF} = 8.75 + 1.5(8.75 - 3.25) = 17$$

أي أنَّ القيم المتطرفة بكبرها هي تلك القيم التي تكون أكبر من 17، ومن ثمَّ لا وجود لقيمة متطرّفة بكبرها في البيانات المقدَّمة.

هكذا نجد أنَّ الشُعيرتان اليمنى واليسرى تبلغان الطول الأعظم لهما حيث لدينا $x_a = \mathbf{LF} = -5$ و $x_\ell = x_e = \mathbf{HF} = 17$، ومن ثمَّ يكون التمثيل الصندوقي للبيانات المُعطاة العرض الآتي:



الشكل [2-17]

٢- لتكن لدينا مجموعة البيانات الآتية:

2, 5, 7, 4, 8, 9, 4, 8, 5, 6, 15, 4

فنجد بعد ترتيبها تصاعدياً أنّ لها العرض الآتي:

2, 4, 4, 4, 5, 5, 6, 7, 8, 8, 9, 15

ومن ثمَّ تكون الأعداد الخمسة لهذه البيانات هي:

$$x_s = 2 \quad \& \quad x_\ell = 15 \quad \& \quad \mathbf{Q}_1 = 4 \quad \& \quad \tilde{x} = \mathbf{Q}_2 = 5.5 \quad \& \quad \mathbf{Q}_3 = 8$$

وأمَّا من أجل تحديد القيم المتطرفة بصغرها فلدينا:

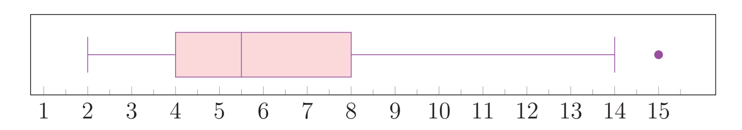
$$\mathbf{LF} = 4 - 1.5(8 - 4) = -2$$

ومن ثمَّ لا وجود لقيم متطرفة بصغرها في البيانات المُعطاة، ولذلك سيتوقف الطرف الأيسر من الشُعيرة اليسرى عند أصغر قيمة للبيانات.

كذلك من أجل تحديد القيم المتطرفة بكبرها لدينا:

$$\mathbf{HF} = 8 + 1.5(8 - 4) = 14$$

ومن ثمَّ لدينا قيمة متطرفة بكبرها في البيانات المقدَّمة وهي $x_\ell = 15$، وهكذا نجد أنَّ الشُعيرة اليمنى تبلغ الطول الأعظم لها، وبالتالي يصبح للتمثيل الصندوقي للبيانات المُعطاة الشكل الآتي.





تمارين



١- قام شخص بشراء أربعةِ أنواعٍ مختلفةٍ من التفاح بأسعار متباينة للكيلو غرام الواحد هي: 5.95، 6.95، 7.95 و6.25. فما هو متوسط سعر الكيلو غرام الواحد من التفاح الذي اشتراه هذا الشخص؟

٢ - بفرض أنَّ 1، 3، 5، 7، 2، 3، 6، 3، 2 و8 تمثِّل أوزان عشرة صناديق من الفاكهة (مقدَّرةً بالكيلوغرام)، فعندئذٍ احسب المتوسط والوسيط، ومن ثمَّ عيّن المنوال (إن وجِد) لهذه البيانات؟

٣- الجدول الآتي يقدِّم لنا درجات خمسين طالباً في أحد المقرّرات الدراسية:

نطاق الدرجات الفئات الفعليّة	أعداد الطلاب الحاصلين على هذه الدرجة التكرار
$49.5 \rightarrow 59.5$	3
$59.5 \rightarrow 69.5$	12
$69.5 \rightarrow 79.5$	10
$79.5 \rightarrow 89.5$	17
$89.5 \rightarrow 99.5$	8
المجموع	**50**

والمطلوب ما يلي:

أ- تعيين الفئات المنوالية إن وجدت وتحديد قيمتها.

ب- تعيين الفئة الوسيطية وتحديد قيمتها.

ج- حساب المتوسط لبيانات هذا الجدول.

٤- إذا كان لدينا خمس مشاهدات بمتوسط $\bar{x} = 10$ ، فما هو مجموع هذه المشاهدات؟

٥- إذا كان لدينا عينة متوسطها $\bar{x} = 10$ ومجموع بياناتها 100، فما هو عدد البيانات في هذه العينة؟

٦- إذا كان لدينا مشاهدات x_1 ، x_2 ، x_3 و x_4 متوسطها $\bar{x} = 6$ ، فما هو متوسط المشاهدات:

$$\frac{x_1}{4} + 0.5 \ ، \ \frac{x_2}{4} + 0.5 \ ، \ \frac{x_3}{4} + 0.5 \ و \ \frac{x_4}{4} + 0.5$$

٧- ما هو مقياس النزعة المركزيّة المناسب لكل من البيانات الآتية، ولماذا؟

a) 2, 4, 4, 7, 7, 7, ?, 8, 10, 13, 15, 19, 21

b) 1, 4, 5, 7, ?, 8, 13, 13, 15, ?, 22, 25, ?

c) 20, 24, 31, 27, 28, 30, 23, 25

d) 20, 23, 21, 121, 28, 30, -20, 28

e) 20, 21, 21, 37, 18, -25, 23, 35

f) 4, 5, 7, 9, ?, ?, 11, 15, 18, 23

علماً أنَّ ? تشير إلى القيم المفقودة في البيانات.

٨- ما هو المنوال (في حال وجوده) لكل من مجموعات المشاهدات الآتية:

أ- مجموعة المشاهدات 3 ،1 ،3 ،5 ،7 ،11 ،6 ،9 و8

ب- مجموعة المشاهدات 7 ،5 ،1 ،6 ،11 ،17 ،12 و21

ج- مجموعة المشاهدات 3 ،3 ،3 ،3 ،3 ،3 و3

د- مجموعة المشاهدات 1 ،3 ،5 ،1 ،3 ،5 ،1 ،3 و5

٩- أي مقياسٍ من مقاييس النزعة المركزيّة يمكن استخدامه مع فصائل الدم O، A، B وAB لمجموعة من الأشخاص.

١٠- سجَّل أحد الطلاب في خمسة مقرّرات لفصل دراسي، وكانت الدرجات التي حصل عليها في كل مقرَّر والأوزان الموافقة لها مقدَّمة في الجدول الآتي.

المقرَّر	O	N	M	L	K
الدرجة	95	73	84	94	85
الوزن	1	2	3	3	2

والمطلوب حساب المعدَّل التراكمي (المتوسط الموزون) للطالب في ذلك الفصل؟

١١- لتكن لدينا البينات الآتية لعينة عشوائية:

4, 7, 5, 4, 8, 5, 17, 8, 7, 13, 2, 4

والمطلوب ما يلي:

أ- تعيين المنوال (أو المناويل في حال وجودها)، ومن ثمَّ حساب المتوسط وبعد ذلك الوسيط أيضاً.

ب- احسب الربيعي الأول والثالث لهذه البيانات.

ج- احسب المئين الـــ 65 لهذه البيانات.

د- بيّن فيما إذا كانت توجد قيم متطرفة في البيانات المُعطاة أم لا.

ه- تقديم الرسم الصندوقي لهذه البيانات موضِّحاً عليه جميع المسمّيات والرموز.

و- بناءً على شكل المخطط الصندوقي للبيانات. هل تلاحظ وجود التواء في توزيع البيانات وإلى أية جهة؟

١٢- لدى معاينة أوزان 20 طفلاً أعمارهم دون العام الواحد وجدنا البيانات الآتية (مقدَّرةً بالكيلو غرام):

10.2	7.7	11.5	7.0
8.8	7.7	10.6	7.3
6.5	7.8	9.8	8.5
9.2	7.7	9.8	8.7
7.6	6.6	9.5	8.6

والمطلوب ما يلي:

أ- تعيين الربيعيَّات الثلاثة Q_1 و Q_2 و Q_3.

ب- تعيين القيم المتطرفة إن وجدت.

ج- تقديم الرسم الصندوقي لهذه البيانات موضِّحاً عليه جميع المسمّيات والرموز.

د- بناءً على شكل المخطط الصندوقي للبيانات. هل تلاحظ وجود التواء في توزيع البيانات وإلى أية جهة؟

١٣- تقدَّم لنا البيانات الآتية الطول لثلاثين شخصاً بالغاً.

178	181	120	171	172	169	158	177	147	172
173	149	158	195	189	163	160	184	201	169
172	173	171	147	178	173	179	218	169	185

والمطلوب ما يلي:

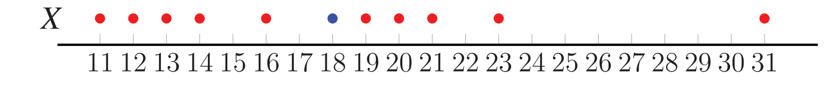
أ- تعيين الربيعيَّات الثلاثة Q_1 و Q_2 و Q_3.

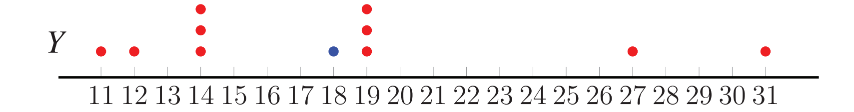
ب- تعيين القيم المتطرفة إن وجدت.

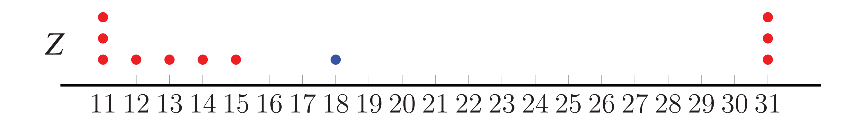
ج- تقديم الرسم الصندوقي لهذه البيانات موضِّحاً عليه جميع المسمّيات والرموز.

د- بناءً على شكل المخطط الصندوقي للبيانات. هل تلاحظ وجود التواء في توزيع البيانات وإلى أية جهة؟

١٤- لتكن لدينا مجموعات البيانات الموضَّحة في العروض الآتية:







والمطلوب ما يلي:

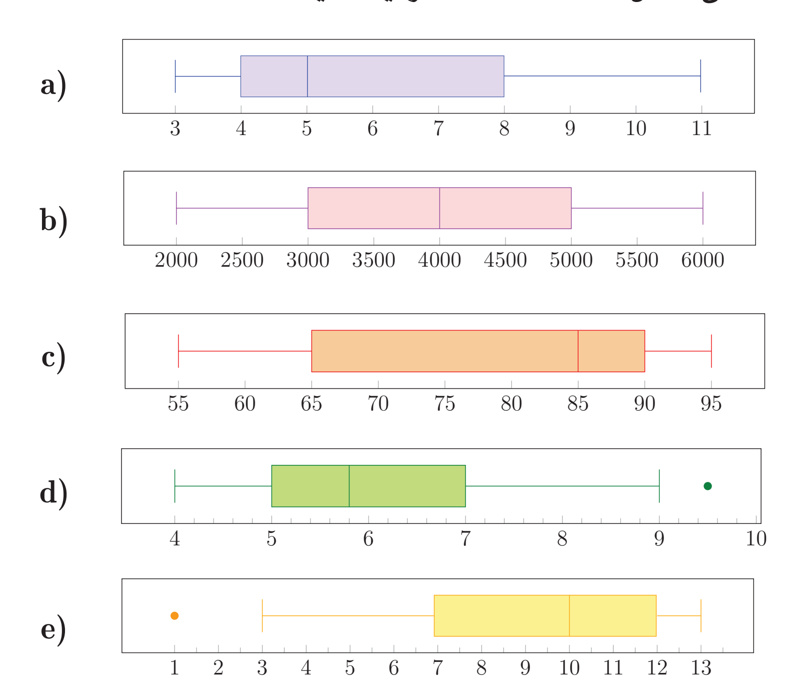
أ- ماذا تمثِّل النقطة الزرقاء على الرسم المقدَّم؟

ب- أي من هذه البيانات تملك منوالاً (أو مناويل)، ثمَّ عيّنها في حال وجودها؟

ج- أيها أكثر تبعثراً حول المتوسط، ولماذا؟

د- عيّن موضع الوسيط لكل منها على الرسم.

١٥- لدينا مجموعات لبيانات مقدَّمة من خلال المخططات الصندوقية الآتية:



والمطلوب ما يلي:

أ- عيِّن الأعداد الخمسة الخاصة بكل مخطط من هذه المخططات.

ب- أي من هذه الأشكال يدلّ على توزيع متناظر؟

ج- أي من هذه الأشكال يدلّ على توزيع ملتوٍ نحو اليمين؟

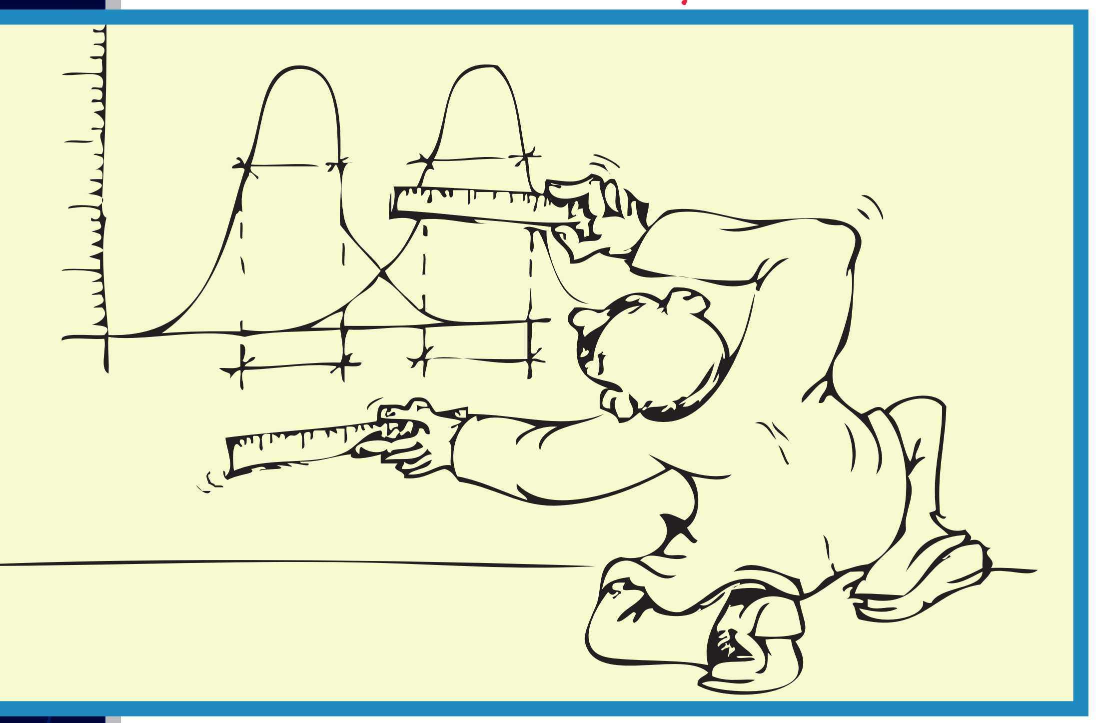
د- أي من هذه الأشكال يدلّ على توزيع ملتوٍ نحو اليسار؟

هـ- أي من هذه الأشكال يشير إلى وجود قيم متطرفة بكبرها؟

و- أي من هذه الأشكال يشير إلى وجود قيم متطرفة بصغرها؟

الفصل الثالث

مقاييس الاختلاف للبيانات
Variability Measures of Data



المقدمة:

لقد ذكرنا في الفصل السّابق أنَّه من المقاييس الضّروريّة للتعرّف على سلوك البيانات الإحصائيّة هي تلك المقاييس التي تهتمّ بتبعثر البيانات حول مقياس نزعتها المركزية (وعلى وجه الخصوص حول متوسطها)، وذلك لأنَّه قد يكون لدينا مجموعتي بيانات أو أكثر من ذات الطبيعة (لها وحدة القياس نفسها) ولها قيمة مقياس النزعة المركزّة نفسه، ولكن تبعثّر البيانات حول مقاييس نزعتها المركزيّة مختلف من مجموعة بيانات إلى أخرى، وبالتالي استخدام مقياس النزعة المركزيّة من أجل المقارنة بين سلوك هذه المجموعات من البيانات يصبح غير ذي جدوى. هذا من جانب، ومن جانبٍ آخر فقد يكون لقيم البيانات في تلك المجموعات وحدات قياس مختلفة، ومن ثمَّ عمليّة المقارنة بين سلوك هذه المجموعات من البيانات يصبح غير ذي جدوى أيضاً حتى لو علمنا كيفية تبعثر البيانات حول مقياس نزعتها المركزيّة لتلك المجموعات من البيانات. لذلك كان لا بدَّ من تقديم مقاييس تساعدنا في مقارنة سلوك البيانات حتى في حال اختلاف وحدات القياس لبيانات تلك المجموعات من البيانات. إنَّ المقاييس التي تساعدنا في مقارنة سلوك البيانات حتى في حال اختلاف وحدات القياس تندرج تحت اسم « مقاييس الاختلاف للبيانات».

٣ - ١
مقاييس التشتت

كما ذكرنا في مقدِّمة هذا الفصل أنَّه من الممكن أن يكون لدينا مجموعتي بيانات أو أكثر من ذات الطبيعة (لها وحدة القياس نفسها) ولها قيمة مقياس النزعة المركزّة نفسه، ولكن تبعثّر البيانات حول مقاييس نزعتها المركزيّة مختلف من مجموعة بيانات إلى أخرى، ولتوضيح ذلك سنأخذ المثال الآتي.

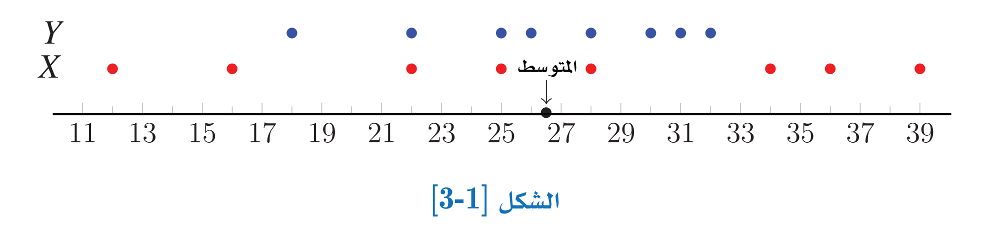
٣-١-١- مثال

لنأخذ مجموعتي البيانات الآتيتين اللتين تمثّلان درجات الحرارة في مدينتين X وY، علماً أنَّ قياس الحرارة (مقدَّرةً بالدرجات المئوية Celsius) قد بدأ الساعة السادسة صباحاً، وأُخِذ قياساً كل ثلاث ساعاتٍ ليومٍ كامل.

الجدول [1–3]

الوقت	3	24	21	18	15	12	9	6
المدينة Y	18	26	28	30	31	32	25	22
المدينة X	16	22	28	34	36	39	25	12

فنجد أنَّ متوسط درجات الحرارة في كلا المدينتين خلال هذا اليوم يساوي 26.5 إلا أنَّ تبعثر قيم درجات الحرارة حول متوسطها يختلف من مدينة إلى أخرى حيث نلاحظ أنَّ بيانات Y قريبة من المتوسط على خلاف بيانات X التي تتناثر مبتعدةً عن المتوسط (انظر الشكل الآتي).



الشكل [3-1]

ممّا سبق يتبيّن لنا ضرورة وجود مقاييس تُحدّد لنا بدقة كيفية تبعثر البيانات حول مقياس نزعتها المركزيَّة ليتكوَّن لدينا انطباعاً أكثر وضوحاً حول سلوك البيانات.

إنَّ المقاييس التي تهتمّ بتحديد مقدار كمّي لقياس مدى تبعثر البيانات حول مقياس نزعتها المركزية تُدعى **مقاييس التّشتت** Dispersion Measures، وسنقدِّم فيما يلي بعضاً منها.

٣-١-٢- الانحراف المعيّاري Standard Deviation

لقد لوحظ أنَّه يمكن النظر إلى قيم الفروقات بين البيانات ومتوسطها كمقياس للتّشتت، ولكن نعلم أنَّ المجموع الجبري لهذه الفروق يساوي الصّفر دوماً، ولذلك طُرِحَت فكرة استخدام مربّعات قيم تلك الفروقات كمقياس للتّشتت، فكانت فكرة تقديم الانحراف المعيّاري الذي سنمهِّد له من خلال التعريف الآتي.

٣-١-٢-١- تعريف (التّباين) (Variance)

لتكن لدينا بيانات عيّنة، فعندئذٍ لتعريف **التّباين** لهذه البيانات (ويُرمز له بـ S^2) سنميِّز بين الحالتين الآتيتين:

١- إذا كانت البيانات خام من قبيل x_1 و x_2 و... و x_n بمتوسط $\bar{x}$، فإنَّ **التّباين** يُعطى من خلال العلاقة الآتية:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 \qquad [3\text{–}1\text{-a}]$$

وبالتعويض في العلاقة السابقة عن قيمة المتوسط $\bar{x}$ بما يساويها بدلالة المجموع وعدد البيانات، ومن ثمَّ إجراء بعض التعديلات على العلاقة يصبح للعلاقة السابقة العرض الآتي:

$$S^2 = \frac{1}{n(n-1)}\left[n\sum_{i=1}^{n}x_i^2 - \left(\sum_{i=1}^{n}x_i\right)^2\right] \qquad [3\text{–}1\text{-b}]$$

وهذه العلاقة الأخيرة تستخدم قيم البيانات مباشرةً دون اللجوء إلى حساب المتوسط أو معرفة قيمته.

٢- إذا كانت البيانات مجمّعة في جدول توزيع تكراري بـ k فئة كما في الجدول [1-2]، وكان متوسطها $\bar{x}$، فإنَّ قيمة **التّباين** تُعطى من خلال العلاقة الآتية:

$$S^2 = \frac{1}{\left(\sum_{i=1}^{k}f_i\right)-1}\sum_{i=1}^{k}f_i\left(x_i - \bar{x}\right)^2 \qquad [3\text{–}2]$$

علماً أنَّ x_i و f_i هما مركز وتكرار الفئة i على الترتيب.

٣-١-٢-٢- تعريف (الانحراف المعيّاري) (Standard Deviation)

يُعرف **الانحراف المعيّاري** على أنَّه الجذر التّربيعي الموجب للتباين، ويُرمز له بـ S، أي أنَّه لدينا:

$$S := +\sqrt{S^2} \qquad [3\text{–}3]$$

وتجدر الإشارة هنا إلى أنَّ **الانحراف المعيّاري** يَستخدم وحدة قياس البيانات نفسها.

٣-١-٢-٣- ملاحظات

١- إنَّ استخدام رمز التّربيع فوق الرّمز S للدّلالة على أنَّ التّباين هو مقدار غير سالب، وأنَّ القيمة النّاتجة عنه تقرأ بالوحدة المربّعة لوحدة القياس المستخدمة في البيانات، فعلى سبيل المثال إذا كانت وحدة القياس للبيانات هي الكيلو غرام فإنَّ قيمة التّباين لهذه البيانات تقرأ بالكيلو غرام المربّع، ولهذا السبب لا يستخدم التّباين بحد ذاته كمقياس للتّشتت.

٢- إنَّ مجموع مربعات انحرافات قيم البيانات x_1 و x_2 و... و x_n عن أي عدد حقيقي a يكون أصغرياً عندما يكون $a = \bar{x}$، ولهذا السبب يُنظر إلى **الانحراف المعيّاري** (**النّاتج عن مفهوم التّباين**) على أنَّه أفضل مقياس للتّشتت.

٣-١-٢-٤- أمثلة

١- لتكن لدينا البيانات الآتية التي تمثِّل درجات 10 طلاب في اختبارٍ قصيرٍ (**من 10 درجات**):

7, 4, 9, 7, 8, 10, 9, 3, 7, 6

ولنقم بحساب الانحراف المعيّاري لدرجات الطلاب.

من أجل تبسيط عملية الحساب وكسب الوقت في الإجابة على مثل هذه المسائل يُفضَّل بناء جدول حسابات على النحو الآتي إذا كنّا سنستخدم المتوسط في حساب الانحراف المعيّاري:

الجدول [2–3]

i	x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1	7	0	0
2	4	-3	9
3	9	2	4
4	7	0	0
5	8	1	1
6	10	3	9
7	9	2	4
8	3	-4	16
9	7	0	0
10	6	-1	1
Total	**70**	**0**	**44**

فنجد أنَّ متوسط درجات الطّلاب $\bar{x}$ يساوي:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{7+4+9+\cdots+7+6}{10} = \frac{70}{10} = 7$$

ومن أجل **الانحراف المعيّاري** لدرجات الطلاب علينا القيام أولاً بحساب التّباين لبيانات درجات الطّلاب فنجد باستخدام العلاقة [3-1-a] أنَّ:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$$
$$= \frac{(7-7)^2 + (4-7)^2 + \ldots + (7-7)^2 + (6-7)^2}{9} = \frac{44}{9} = 4.\bar{8}$$

ومنه ينتج لدينا أنَّ الانحراف المعيّاري لدرجات الطّلاب يساوي:

$$S = +\sqrt{S^2} = +\sqrt{4.9} = 2.21$$

٢- عندما نتسوّق في أحد المتاجر للمواد التّموينيّة لاحظنا أنّه كُتِبَ على إحدى السلع عبارة "الوزن الصافي ... كغ $\pm \cdots$ غ"، وللتّحقّق من ذلك قمنا بوزن 15 قطعة من هذه السّلعة فوجدنا القيم الآتية (مقدَّرةً بالكيلو غرام):

4.87, 5.2, 5.1, 4.95, 4.85, 4.85, 4.75, 4.97, 5.15, 5.05, 5.2, 5.05, 5.12, 4.98, 4.91

ولنقم بتعيين قيمة الوزن الصافي لهذه السلعة مع تحديد قيمة $\pm \cdots$ التي كتبت إلى جانبها على هذه السلعة.

الجواب: إنَّ قيمة الوزن الصّافي لهذه السّلعة يقصد به متوسط الأوزان لجميع قطع هذه السّلعة، وأمَّا العبارة $\pm \cdots$ غ فيقصد بها أنَّ معظم هذه السّلع يمكن لها أن تقل أو تزيد عن المتوسط بالمقدار المشار إليه وهو في الواقع قيمة الانحراف المعيّاري لأوزان هذه السّلعة، وبحسب البيانات المُعطاة يكون لدينا:

الجدول [3-3]

i	x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1	4.87	-0.13	0.0169
2	5.20	0.2	0.0400
3	5.10	0.1	0.0100
4	4.95	-0.05	0.0025
5	4.85	-0.15	0.0225
6	4.85	-0.15	0.0225
7	4.75	-0.25	0.0625
8	4.97	-0.03	0.0009

i	x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
9	5.15	0.15	0.0225
10	5.05	0.05	0.0025
11	5.20	0.2	0.0400
12	5.05	0.05	0.0025
13	5.12	0.12	0.0144
14	4.98	-0.02	0.0004
15	4.91	-0.09	0.0081
Total	**75**	**0**	**0.2682**

ومنه نجد أنَّ متوسط الوزن لهذه السلعة يساوي:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{4.87 + 5.2 + \cdots + 4.98 + 4.91}{15} = \frac{75}{15} = 5\,\text{Kg}$$

والآن لحساب الانحراف المعيّاري يجب علينا حساب التّباين أولاً، وبما أنَّنا قمنا آنفاً بحساب المتوسط للبيانات فإنَّه باستخدام العلاقة [3–1-a] يكون لدينا:

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2 = \frac{(4.87-5)^2 + \cdots + (4.91-5)^2}{14} = \frac{0.2682}{14} = 0.01916$$

ومن ثمَّ ينتج أنَّ قيمة الانحراف المعيّاري لأوزان هذه السّلعة يساوي:

$$S = +\sqrt{S^2} = +\sqrt{0.01916} = 0.138 \quad \text{Kg}$$

وبتحويلها للغرام يصبح للعبارة "الوزن الصّافي ... كغ ± ··· غ" العرض الآتي:

الوزن الصافي 5 كغ 138 ± غ

٣- بالعودة إلى بيانات المثال (٣-١-١) لنقم بحساب الانحراف المعيّاري لكل من مجموعتي البيانات X و Y مستخدمين في ذلك العلاقة [3–1-b].

لنقم أولاً بحساب التّباين لكل من مجموعتي البيانات X و Y، ومن أجل ذلك يُفضَّل بناء جدول حسابات على النّحو الآتي:

الجدول [3-4]

i	x_i	x_i^2	y_i	y_i^2
1	12	144	22	484
2	25	625	25	625
3	39	1521	32	1024
4	36	1296	31	961
5	34	1156	30	900
6	28	784	28	784
7	22	484	26	676
8	16	256	18	324
Total	**212**	**6266**	**212**	**5778**

فنجد من أجل مجموعة البيانات X أنَّ التّباين يساوي:

$$S_X^2 = \frac{1}{n(n-1)}\left[n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right] = \frac{1}{8(8-1)}\left[8(6266) - (212)^2\right] = 92.57$$

ومنه نجد أنَّ قيمة الانحراف المعيّاري للبيانات X هي:

$$S_X = +\sqrt{S_X^2} = +\sqrt{92.57} = 9.62$$

وأمّا من أجل مجموعة البيانات Y لدينا التّباين يساوي:

$$S_Y^2 = \frac{1}{n(n-1)}\left[n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2\right] = \frac{1}{8(8-1)}\left[8(5778) - (212)^2\right] = 22.86$$

ومنه نجد أنَّ قيمة الانحراف المعيّاري للبيانات Y هي:

$$S_Y = +\sqrt{S_Y^2} = +\sqrt{22.86} = 4.78$$

وهكذا نجد أنَّ قيم درجات الحرارة في المدينة X تتبعثر حول متوسطها بشكلٍ أكبر ممّا تتبعثر فيه قيم درجات حرارة المدينة Y حول متوسطها.

٤- الجدول الآتي يقدِّم عدد الحوادث المروريّة خلال شهرٍ معيّنٍ لكلِّ فئةٍ عمريّةٍ في مدينة ما X.

الجدول [3-5]

الفئة العمرية	عدد الحوادث
$18 \to 24$	37
$24 \to 30$	28
$30 \to 36$	17
$36 \to 42$	8
$42 \to 48$	5
Total	**95**

ولنقم بحساب متوسط عدد الحوادث في هذه المدينة، ومن ثمَّ حساب الانحراف المعيّاري المرافق له.

الجواب: من أجل ذلك لنقدّم جدول الحسابات الآتي:

الجدول [3-6]

i	التكرارات f_i	مراكز الفئات x_i	$f_i \cdot x_i$	$x_i - \bar{x}$	$f_i(x_i - \bar{x})^2$
1	37	21	777	-6.695	1658.452
2	28	27	756	-0.695	13.5247
3	17	33	561	5.305	478.4314
4	8	39	312	11.305	1022.424
5	5	45	225	17.305	1497.315
Total	**95**	----------------	**2631**	-----------	**4670.147**

والآن لنحسب المتوسط، فنجد:

$$\bar{x} = \frac{1}{\sum_{i=1}^{k} f_i} \sum_{i=1}^{k} f_i \cdot x_i = \frac{(37 \times 21) + \cdots + (5 \times 45)}{95} = \frac{2631}{95} = 27.695$$

وأمّا لحساب التّباين فإنَّنا سنستخدم العلاقة [3-2]، فيكون لدينا:

$$S^2 = \frac{1}{\left(\sum_{i=1}^{k} f_i\right) - 1} \sum_{i=1}^{k} f_i \left(x_i - \bar{x}\right)^2$$

$$= \frac{37\,(21 - 27.695)^2 + \ldots + 5\,(45 - 27.695)^2}{94} = \frac{4670.147}{94} = 49.68$$

وبالتالي فإنَّ قيمة الانحراف المعيّاري تساوي:

$$S = +\sqrt{S^2} = +\sqrt{49.68} = 7.05$$

٣-١-٢-٥- مزايا وعيوب الانحراف المعيّاري

١- إنَّ الانحراف المعيّاري هو أفضل مقاييس التّشتت بلا منازع كما ذكرنا ذلك سابقاً، ويأخذ بالحسبان جميع قيم البيانات.

٢- إذا وجِدَتْ قيم متطرّفة فإنَّ المتوسط سيتأثر بها، ومن ثمَّ سينتقل هذا التأثير على قيمة الانحراف المعيّاري أيضاً، بمعنى أنَّ الانحراف المعيّاري يتأثر بالقيم المتطرفة أيضاً، وهذه تُعد من سلبيات الانحراف المعيّاري كمقياس للتّشتت.

٣- إذا فُقِدتْ إحدى أو بعض البيانات فعندئذٍ يُصبح الانحراف المعيّاري عديم الفائدة، وهذه تعد سلبيّة أخرى للانحراف المعيّاري أيضاً.

هكذا نلاحظ أنَّه في حال فقدان بعض قيم البيانات يجب علينا البحث عن مقياس آخر للتّشتت، وفي حال وجود قيم متطرفة يُفضَّل استخدام مقاييس أخرى للتّشتت أيضاً.

إنَّ المقياس الآتي يقدِّم لنا بعض الحلول الجزئيّة للمشكلات التي ذكرناها سابقاً.

٣-١-٢- المدى Range

لقد قمنا في الفصل الأول بتقديم تعريف المدى لمجموعة من البيانات الخام (التعريف ١-٣-١) من خلال العلاقة [2-1]، وأمَّا إذا كانت البيانات مجمَّعة في جدول توزيع تكراري بـ k فئة، فإنَّ قيمة **المدى** لبيانات هذا الجدول تُعطى من خلال العلاقة الآتية:

$$\mathbf{R} = x_k - x_1 \qquad [3\text{-}4]$$

علماً أنَّ x_1 و x_k هما مركز الفئة الأولى والأخيرة على الترتيب.

في الواقع يُعدّ المدى مقياساً للتّشتت في حال فقدان بعض البيانات غير الواقعة على أطراف البيانات بعد ترتيبها، أي أنَّه عندما تكون أكبر وأصغر قيمة في البيانات ليست في عداد البيانات المفقودة. لكن لا يُنظر إلى هذا المقياس على أنَّه مقياسٌ جيدٌ للتّشتت رغم أنَّه يُعطي صورةً عن مدى تشـتت مجموعة من البيانات، إذ إنَّه وفي كثيـر من الحالات (وخاصة لدى العينات كبيرة الحجم) لا يُظهر لنا بوضوحٍ توضَّع البيانات حول متوسطها لأنَّه يعتمد على الفرق بين أكبر وأصغر قيمةٍ فقط، والمثال الآتي يوضِّح لنا ذلك.

٣-١-٢-١- مثال

لنأخذ مجموعتي البيانات الآتيتين:

الجدول [3–7]

مجموعة البيانات X	3	8	5	10	7	15	7	9	9
مجموعة البيانات Y	7	8	5	8	17	6	8	5	5

فنجد أنَّ لكل من مجموعتي البيانات X وY المتوسط نفسه $\bar{x} = \bar{y} = 8$ والمدى نفسه أيضاً:

$$\mathbf{R_X} = 15 - 3 = 12 \qquad \& \qquad \mathbf{R_Y} = 17 - 5 = 12$$

كيفية انتشار البيانات حول المتوسط لكل من مجموعتي البيانات X وY

وأمَّا لحساب الانحراف المعيّاري فلدينا:

الجدول [3-8]

i	x_i	$(x_i - \bar{x})^2$	y_i	$(y_i - \bar{y})^2$
1	3	25	7	1
2	8	0	8	0
3	5	9	5	9
4	10	4	8	0
5	7	1	17	81
6	15	49	6	4
7	7	1	8	0
8	9	1	5	9
Total	**64**	**90**	**64**	**104**

ومن ثمَّ نجد التّباين لمجموعة البيانات X وY على الترتيب يساوي:

$$S_X = \sqrt{\frac{\sum_{i=1}^{8}(x_i - \bar{x})^2}{7}} = \sqrt{\frac{90}{7}} = 3.5857$$

$$S_Y = \sqrt{\frac{\sum_{i=1}^{8}(y_i - \bar{y})^2}{7}} = \sqrt{\frac{104}{7}} = 3.8545$$

وهكذا نلاحظ وجودَ تفاوتٍ واضحٍ في تبعثر البيانات حول متوسطها لكلٍّ من مجموعتي البيانات.

٣-١-٢-١- مزايا وعيوب المدى كمقياس للتَّشتت

بالرغم من أنَّ المدى يعدّ مقياساً ضعيفاً للتّشتت إلا أنَّه يتمتع بمجموعة من المزايا تجعل منه مقياساً مرغوباً في بعض الحالات، فمن مزايا المدى:

١- إنَّه بسيط جداً وسهل الحساب.

٢- يكثر استخدامه عند الإعلان عن حالات المُناخ: كدرجات الحرارة، الرطوبة والضغط الجوي.

٣- يستخدم في بعض مجالات مراقبة الجودة.

من عيوب المدى:

١- يعتمد على قيمتين فقط، ولا يأخذ جميع القيم في الحسبان، ومن ثمَّ تكون قيمته أقل دقةً من المقياس السابق.

٢- يتأثر بالقيم المتطرفة بشكل كبير جدّاً لأنَّه في الأصل يعتمد على القيم التي تقع على الأطراف.

إذاً فلا بدَّ من البحث عن مقياس آخر يتجاوز السلبيّات للمقياس السّابق، وفي هذا الإطار نجد أنَّ مقياس التّشتت الآتي يقدّم لنا حلاً جزئياً آخر لمشكلة وجود القيم المتطرّفة في البيانات.

٣-١-٣- تعريف (المدى الرُبَيْعي) (Interquartile Range)

لتكن x_1 و x_2 و... و x_n بيانات خام لعيِّنة مُعطاة، فعندئذٍ يُعرَّف **المدى الربيعي** (ويُرمز له **IQR**) من خلال العلاقة الآتية:

$$IQR = Q_3 - Q_1 \qquad [3-5]$$

علماً أنَّ Q_1 و Q_3 هما الربيعي الأول والثّالث للبيانات على الترتيب.

لاحظ أنَّ هذا المقياس يَحسب المدى لنصف عدد البيانات الواقعة في الوسط بعد ترتيبها تصاعدياً، ومن ثمَّ فإنَّ القيم المتطرفة ستصبح خارج نطاق هذا المقياس. من جهة أخرى، فعلى الرغم من أنَّ هذا المقياس أفضل من المدى إلاَّ أنَّه يعتمد في قراره على قيمتين من البيانات فقط ولا يأخذ في الحسبان مواضع جميع البيانات، ولهذا السبب يُنظر إلى هذا المقياس على أنَّه مقياس ضعيف للتّشتت أيضاً، ولكنَّه يُعدّ مقبولاً في حال فقدان بعض البيانات المعلوم ترتيبها وغير الموافقة لقيمتي الرُبيْعيين الأول والثالث.

أخيراً نشير إلى أنَّ نصف قيمة المدى الربيعي $Q = IQR/2$ تدعى **الانحراف الربيعي** Quartile Deviation، وتستخدم كمقياس للتّشتت أيضاً.

٣-١-٣-١- أمثلة

١- لتكن لدينا البيانات الإحصائية الآتية:

8, 5, 3, 11, 5, 4, 7, 5, 3, 9, 6

ولنقم بحساب المدى الربيعي لهذه البيانات.

الجواب: من أجل ذلك يجب حساب الربيعي الأول والثالث، وهذا يتطلب ترتيب البيانات أولاً حيث لدينا:

البيانات بعد ترتيبها	11	9	8	7	6	5	5	5	4	3	3
رموز البيانات بعد ترتيبها	x_{11}	x_{10}	x_9	x_8	x_7	x_6	x_5	x_4	x_3	x_2	x_1

وبما أنَّ $n = 11$ فإنَّ رتبة الرُبيْعي Q_1 و Q_3 ستكون على الترتيب هي:

$$q_1 = \frac{n+1}{4} = \frac{11+1}{4} = 3 \qquad \& \qquad q_3 = \frac{3(n+1)}{4} = \frac{3(11+1)}{4} = 9$$

ومن ثمَّ يكون لدينا:

$$Q_1 = x_k + s(x_{k+1} - x_k) = x_3 + 0(x_4 - x_3) = 4$$

$$Q_3 = x_k + s(x_{k+1} - x_k) = x_9 + 0(x_{10} - x_9) = 8$$

وهكذا يكون المدى الربيعي لمجموعة البيانات هو:

$$IQR = Q_3 - Q_1 = 8 - 4 = 4$$

٢- بالرجوع إلى بيانات الجدول [3-7] حيث لدينا $n = 8$ فإنَّنا نجد ما يلي:

أ- من أجل البيانات X: نجد للبيانات بعد ترتيبها تصاعدياً العرض الآتي:

البيانات X بعد ترتيبها	15	10	9	8	7	7	5	3
رموز البيانات بعد ترتيبها	x_8	x_7	x_6	x_5	x_4	x_3	x_2	x_1

ومن ثمَّ يكون لرتبة الرُبيْعي Q_1 و Q_3 القيمتين الآتيتين على الترتيب:

$$q_1 = \frac{n+1}{4} = \frac{8+1}{4} = 2.25 \qquad \& \qquad q_3 = \frac{3(n+1)}{4} = \frac{3(8+1)}{4} = 6.75$$

ومنه يكون لدينا:

$$\mathbf{Q}_1 = x_k + s(x_{k+1} - x_k) = x_2 + 0.25(x_3 - x_2) = 5 + 0.5 = 5.5$$

$$\mathbf{Q}_3 = x_k + s(x_{k+1} - x_k) = x_6 + 0.75(x_7 - x_6) = 9 + 0.75 = 9.75$$

وهكذا يكون المدى الربيعي لمجموعة البيانات هو:

$$\mathbf{IQR} = \mathbf{Q}_3 - \mathbf{Q}_1 = 9.75 - 5.5 = 4.25$$

ب- من أجل البيانات Y: فنجد للبيانات الترتيب التصاعدي الآتي:

البيانات Y بعد ترتيبها	17	8	8	8	7	6	5	5
رموز البيانات بعد ترتيبها	x_8	x_7	x_6	x_5	x_4	x_3	x_2	x_1

ومن ثمَّ يكون لرتبة الرُبيْعي $\mathbf{Q}_1$ و $\mathbf{Q}_3$ القيمتين الآتيتين على الترتيب:

$$q_1 = \frac{n+1}{4} = \frac{8+1}{4} = 2.25 \qquad \& \qquad q_3 = \frac{3(n+1)}{4} = \frac{3(8+1)}{4} = 6.75$$

ومنه يكون لدينا:

$$\mathbf{Q}_1 = x_k + s(x_{k+1} - x_k) = x_2 + 0.25(x_3 - x_2) = 5 + 0.25 = 5.25$$

$$\mathbf{Q}_3 = x_k + s(x_{k+1} - x_k) = x_6 + 0.75(x_7 - x_6) = 8 + 0 = 8$$

وهكذا يكون المدى الربيعي لمجموعة البيانات هو:

$$\mathbf{IQR} = \mathbf{Q}_3 - \mathbf{Q}_1 = 8 - 5.25 = 2.75$$

▶

بذلك نجد أنَّ مجموعة البيانات X تتبعثر حول متوسطها بشكلٍ أكبر من تبعثر مجموعة البيانات Y حول متوسطها، وهكذا نلاحظ أنَّ ما عَجِزَ عنه المدى (في إظهار فروق التّشتت لمجموعتي البيانات) أنجزه المدى الربيعي.

٣ - ٢
مُعاملات من أجل مقارنة التشتت لمجموعتي بيانات أو أكثر

لقد لاحظنا أنَّ مقاييس التّشتت تعتمد على وحدة القياس المستخدمة من أجل البيانات، ولذلك يصعب علينا إجراء المقارنة بين تشتت مجموعتي بيانات أو أكثر عندما تكون وحدات القياس المستخدمة من أجل كل منها مختلفة عن الأخرى، ولذلك كان لابد من وضع معيَّار يُمكّننا من الحكم على مثل هذه المقارنات حتى في حال كانت وحدات القياس مختلفة بعضها عن البعض الآخر. يقدِّم لنا المعيَّار الآتي حلاً لهذه الإشكاليّة.

٣-٢-١- مُعامل التَغيرُّ (أو الاختلاف) Coefficient of Variation

لتكن لدينا مجموعة بيانات عيّنة بمتوسط $\bar{x} \neq 0$ وانحراف معيَّاري S، فعندئذٍ يُعرَّف **مُعامل التَغيرُّ** (والذي يُرمز له بـ **CV**) لهذه البيانات من خلال العلاقة الآتية:

$$\mathbf{CV} := \frac{S}{\bar{x}} \times 100\ \% \qquad [3\text{–}6]$$

ويقرأ النّاتج كنسبةٍ مئويّةٍ كما هو واضح.

لاحظ أنّ مُعامل التغيرُّ يقدِّم لنا قيمة تجعلنا نشعر بمدى التغيرّ الحاصل لمتغيرٍّ ما، ويفيدنا في مقارنة إحصائية لدرجة التّباين من سلسلة بيانات إلى أخرى حتى ولو كانت وحدات القياس تختلف بشكل كبير بعضها عن البعض الآخر، وذلك لأنَّ نسبة الانحراف المعيّاري إلى المتوسط الحسابي تلغي خاصيّة وحدة القياس المستعملة (فمثلاً: متر على متر، كيلوغرام على كيلوغرام، فولط على فولط أو .... حيث يختفي أثر وحدة القياس) وإظهار مفعول التّباين بآن واحد، ومن ثمَّ تقديم هذه القيمة كنسبة مئوية خاصة بالبيانات.

أخيراً نشير إلى أنَّ |**CV**| (القيمة المطلقة لمُعامل الاختلاف) تُعرف باسم "الانحراف المعيّاري النسبي" Relative Standard Deviation (**RSD**)، ويُعبرَّ عنها كنسبةٍ مئويّةٍ أيضاً.

◀ ٣-٢-١-١- مثال

لتكن لدينا مجموعتي بيانات تمثِّل الطول والوزن لستة أطفال في سن العاشرة، ومقدَّمتين من خلال الجدول الآتي:

الجدول [3–9]

الشخص	F	E	D	C	B	A
الطول X بـ سم	129	119	137	128	114	123
الوزن Y بـ كغ	33	31	34	29	31	25

فهل تشير هذه البيانات إلى أنَّ تبعثر قياسات الطول حول متوسطها أصغر من تبعثر قياسات الوزن حول متوسطها؟

الحل: نلاحظ أنَّ وحدة القياس بين مجموعتي البيانات مختلفة، وبالتالي لا يمكننا استخدام الانحراف المعيّاري من أجل إجراء المقارنة المطلوبة. لذلك سنستخدم معامل الاختلاف للحصول على القرار.

الآن، بحساب قيم المتوسط والانحراف المعيّاري لكل من مجموعتي البيانات نجد أنَّ:

$$\bar{x} = 125 \ \ cm \qquad \& \qquad S_X = 8.12 \ \ cm$$

$$\bar{y} = 30.5 \ \ Kg \qquad \& \qquad S_Y = 3.21 \ \ Kg$$

وبالتالي تكون قيمة مُعامل التغيُّر لمجموعة البيانات X هي:

$$\mathbf{CV}_X = \frac{S_X}{\bar{x}} \times 100 = \frac{8.12}{125} \times 100 = 6.496 \ \%$$

في حين نجد قيمة مُعامل التغيُّر لمجموعة البيانات Y هي:

$$\mathbf{CV}_Y = \frac{S_Y}{\bar{y}} \times 100 = \frac{3.21}{30.5} \times 100 = 10.525 \ \%$$

وهكذا نجد أنَّ قيمة مُعامل تغيُّر الوزن أكبر من قيمة مُعامل تغيُّر الطول لهذه العيِّنة من الأطفال، ومن ثمَّ تبعثر قياسات الطّول حول متوسطها أقلّ من تبعثر قياسات الوزن حول متوسطها.

▶

الآن، وفي حال تعذر حساب المتوسط أو التّباين لسبب ما (مثل فقدان بعض البيانات) فإنَّه يجب البحث عن مُعامل آخر ينجز لنا عملية المقارنة. إنَّ المعيّار الآتي الذي يُحسب بدلالة الرّبيعيين الأول والثالث يمكن استخدامه بدلاً من مُعامل التغيُّر.

٣-٢-٢- مُعامل التّشتت Coefficient of Dispersion

لتكن لدينا مجموعة بيانات عينة مُعطاة، فعندئذٍ يُعرَّف **مُعامل التّشتت** (والذي يُرمز له بـ **CD**) لبيانات هذه العينة من خلال العلاقة الآتية:

$$\mathbf{CD} := \frac{\mathbf{Q}_3 - \mathbf{Q}_1}{\mathbf{Q}_3 + \mathbf{Q}_1} \times 100 \ \% \qquad [3\text{–}7]$$

علماً أنَّ $\mathbf{Q}_1$ و $\mathbf{Q}_3$ هما الرُّبيْعي الأول والثالث للبيانات على الترتيب.

◀ ٣-٢-٢-١- مثال

تقدَّمت مجموعة من الطّلاب لاختبارين (مقابلة وتحريري) في مُقرَّر دراسي، فكانت لهم النتائج الآتية التي تظهر فقدان بعض الدرجات (المعلوم ترتيبها) في اختبار المقابلة:

الجدول [3–10]

الطالب	J	I	H	G	F	E	D	C	B	A
نتائج الاختبار التحريري X	25	36	37	45	43	49	35	48	42	35
نتائج اختبار المقابلة Y	19	18	17	?	15	13	?	10	8	05

فهل تشير هذه المُعطيات إلى أنَّ درجات الاختبار التحريري تتبعثر حول متوسطها أكثر ممّا هو لدى اختبار المقابلة؟

الحل: بالطبع من كون إحدى مجموعتي البيانات تحتوي على بيانات مفقودة فليس من المنطقي أن نستخدم مُعامل التغيّر من أجل البيانات الكاملة ومُعامل التّشتت من أجل البيانات المنقوصة لكي نعطي قرارنا في عملية المقارنة، وإنّما علينا استخدام المعيّار نفسه للوصول إلى القرار الصحيح. لذلك سنقوم بحساب مُعامل التّشتت لكل من مجموعتي البيانات، ولأجل ذلك سنقوم أولاً بترتيب مجموعة بيانات الاختبار التحريري فقط لأنَّ بيانات اختبار المقابلة قُدِّمَت مرتبةً، فيكون لدينا العرض الآتي:

نتائج الاختبار التحريري بعد ترتيبها	49	48	45	43	42	37	36	35	35	25
رموز القيم بعد ترتيبها	x_{10}	x_9	x_8	x_7	x_6	x_5	x_4	x_3	x_2	x_1

فنجد من أجل هذه البيانات أنَّ لرتبة الرُّبيْعيين $\mathbf{Q}_1$ و $\mathbf{Q}_3$ القيمتين الآتيتين على الترتيب:

$$q_1 = \frac{n+1}{4} = \frac{10+1}{4} = 2.75 \qquad \& \qquad q_3 = \frac{3(n+1)}{4} = \frac{3(10+1)}{4} = 8.25$$

ومن ثمَّ يكون لدينا:

$$\mathbf{Q}_1 = x_2 + 0.75\,(x_3 - x_2) = 35 + 0.75\,(35 - 35) = 35$$

$$\mathbf{Q}_3 = x_8 + 0.25\,(x_9 - x_8) = 45 + 0.25\,(48 - 45) = 45.75$$

وبالتالي تكون قيمة مُعامل التّشتت لبيانات الاختبار التحريري هي:

$$\mathbf{CD}_X = \frac{\mathbf{Q}_3 - \mathbf{Q}_1}{\mathbf{Q}_3 + \mathbf{Q}_1} \times 100 = \frac{45.75 - 35}{45.75 + 35} \times 100 = 13.31\ \%$$

وأمّا من أجل رتبة الرُّبيْعيين $\mathbf{Q}_1$ و $\mathbf{Q}_3$ الخاصّة باختبار المقابلة فنجد لهما القيمتين الآتيتين على الترتيب:

$$q_1 = \frac{n+1}{4} = \frac{10+1}{4} = 2.75 \qquad \& \qquad q_3 = \frac{3(n+1)}{4} = \frac{3(10+1)}{4} = 8.25$$

ومن ثمَّ يكون لدينا:

نتائج اختبار المقابلة (مرتَّبة)	19	18	17	?	15	13	?	10	8	5
رموز القيم المرتَّبة	y_{10}	y_9	y_8	y_7	y_6	y_5	y_4	y_3	y_2	y_1

$$\mathbf{Q}_1 = y_2 + 0.75\,(y_3 - y_2) = 8 + 0.75\,(10 - 8) = 9.5$$

$$Q_3 = y_8 + 0.25\,(y_9 - y_8) = 17 + 0.25\,(18 - 17) = 17.25$$

ومنه نجد أنَّ قيمة مُعامل التّشتت لبيانات اختبار المقابلة تساوي:

$$CD_Y = \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100 = \frac{17.25 - 9.5}{17.25 + 9.5} \times 100 = 28.97\ \%$$

وهكذا ينتج لدينا أنَّ بيانات اختبار المقابلة تتبعثر حول متوسطها بشكل أكبر ممّا هو عليه الحال من أجل بيانات الاختبار التحريري.

٣ - ٣
الدرجة المعيارية Z

الآن، وقبل ختام هذا الفصل سنقدّم مفهوماً متعلّقاً بقيمة المتوسط والانحراف المعيّاري لمجموعة بيانات، ويُدعى **الدرجة المعيّارية Z** (وسنستخدم عبارة الدرجة المعيّارية على سبيل الاختصار).

من فوائد الدرجة المعيّارية أنّها تُعطينا صورةً عن موضع البيان بالنسبة إلى متوسط البيانات، ولذلك يمكننا أن نقارن بين قيمتين لكل منهما موضع نسبيٌّ مختلف في مجموعتي بيانات مختلفتين (وقد يكون لهما قيم مختلفة للمتوسط)، فعلى سبيل المثال يمكننا مقارنة مستوى أداء طالب في جامعة ما مع مستوى أداء طالب آخر من جامعة أخرى، أو مقارنة مستوى أداء طالبٍ في مدرسةٍ حكوميّةٍ مع مستوى أداء طالبٍ آخر من مدرسةٍ أهليّة.

٣-٣-١- تعريف (الدرجة المعيّارية)

لتكن x_1 و x_2 و $\cdots$ و x_n بيانات عيّنة بمتوسط $\bar{x}$ وانحراف معيّاري S، فعندئذٍ **الدرجة المعيّارية** Z لقيمة x_i من هذه البيانات (والتي سنرمز لها بـ z_i) تُعرّف من خلال العلاقة الآتية:

$$z_i := \frac{x_i - \bar{x}}{S} \quad ; i = 1, 2, \ldots , n \qquad [3–7]$$

٣-٣-٢- ملاحظات

١- إنّ العملية التي نُفِّذت على القيمة x_i بوساطة العلاقة [3-7] تُدعى عمليّة استعيّار للقيمة x_i ولذلك سُميَّت z_i بالدرجة المعيّاريّة لـ x_i.

٢- نلاحظ أنّ قيمة هذا المعيّار هي مؤشّر على انحراف القيمة x_i عن المتوسط، ومن ثمَّ فإنّها تُحدِّد موقع x_i من المتوسط اتجاهاً وبعداً، فالاتجاه تُحدّده إشارة (- أو+)، فإذا كانت قيمة z_i موجبة فإنَّ ذلك يعني أنَّ x_i أكبر من المتوسط، والعكس إذا كانت قيمة z_i سالبة. أمَّا البُعد فتعني كِبَرَ القيمة المطلقة لـ z_i، فكلما كَبُرت القيمة المطلقة لـ z_i دلَّ ذلك على ابتعاد القيمة x_i عن المتوسط.

◄ ٣-٣-٣- أمثلة

١- في إطار التقدّم لوظيفة مدرّس لدى وزارة التعليم تقدَّم خريجين جامعيين A وB من كلية X وY على الترتيب في جامعةٍ ما. فإذا علمت أنَّ المعدَّل التراكمي للشخص A يساوي 4.7 في حين كان المعدَّل التراكمي للشخص B يساوي 4.4، فهل يمكننا الادّعاء أنَّ أداء الشخص A أفضل من أداء الشخص B؟

الحل: في الواقع إنَّ الردّ على هذا التساؤل يتطلب معرفة مستوى كل من هذين الشخصين في كليّته لأنَّهما لا يخضعان لتعليم موحَّد سواءً من حيث البيئة المحيطة بالشخص أو من حيث طبيعة العلم الذي يتلقاه، ولذلك الإجابة على السؤال المطروح ليست بهذه البساطة، ويجب أن يُؤخذ في الحسبان المعدَّل التراكمي لزملاء كل واحد منهم، ومن ثمَّ يُبحث في الإجابة بناءً على هذه المُعطيات.

فلو افترضنا أنَّ متوسط المعدَّل التراكمي لخريجي الكلية X يساوي 4.6 بانحراف معيَّاري يساوي 0.3، وأنَّ متوسط المعدَّل التراكمي لخريجي الكلية Y يساوي 4.2 بانحراف معيَّاري يساوي 0.4، فعندئذٍ بحساب الدرجة المعيّارية لكل من هذين المتقدمين نجد الآتي:

$$z_{\mathbf{A}} = \frac{4.7 - 4.6}{0.3} = 0.\overline{3} \qquad \& \qquad z_{\mathbf{B}} = \frac{4.4 - 4.2}{0.4} = 0.5$$

وهكذا نجد أنَّ الدرجة المعيّارية للشخص B أعلى من الدرجة المعيّارية للشخص A بفارقٍ كبيرٍ، وهذا يعني أنَّ أداء الشخص B أفضل من أداء الشخص A رغم أنَّ المعدَّل التراكمي للشخص B أقل من المعدَّل التراكمي للشخص A.

٢- لتكن لدينا مجموعتي البيانات الآتيتين:

a) 7, 9, 8, 6, 2, 7, 9, 7, 3, 6, 5, 3

b) 8, 6, 7, 8, 8, 9, 6, 9, 5, 4

ولنوجد قيمة الدرجة المعيّارية للقيمة $x = 8$ في كلٍّ من مجموعتي البيانات المُعطاة.

الحل: من أجل مجموعة البيانات (a) نجد أنَّ متوسطها $\overline{x} = 6$ وانحرافها المعيّاري $S = 2.335$، ومن ثمَّ تكون الدرجة المعيّارية للقيمة $x = 8$ في مجموعة البيانات (a) تساوي:

$$z_a = \frac{x - \overline{x}}{S} = \frac{8 - 6}{2.335} = 0.857$$

وأمَّا من أجل مجموعة البيانات (b) فنجد أنَّ متوسطها $\overline{x} = 7$ وانحرافها المعيّاري $S = 1.7$، ومن ثمَّ تكون الدرجة المعيّارية للقيمة $x = 8$ في مجموعة هذه البيانات (b) تساوي:

$$z_b = \frac{x - \overline{x}}{S} = \frac{8 - 7}{1.7} = 0.588$$

وهكذا نجد أنَّ الدرجة المعيّارية للقيمة $x = 8$ في مجموعة البيانات (a) أكبر من الدرجة المعيّارية للقيمة نفسها في مجموعة البيانات (b).

٣-٣-٤- ملاحظات

١- لو حُسِبت الدرجة المعيّارية لجميع عناصر العيّنة فعندئذٍ سنلاحظ أنَّ البيانات الناتجة عن هذه الدرجات المعيّارية لها متوسط يساوي الصفر وانحراف معيَّاري يساوي الواحد.

٢- يُؤخذ على الدرجة المعيّارية أنَّها عند قيمة الصفر يكون للبيان الذي استعيرِّت قيمته درجة قيمة المتوسط $\bar{x}$، ولذلك قد لا يكون معناها واضحاً لدى الكثير، وعلاوةً على ذلك فقد تكون هذه الدرجة المعيّارية سالبةً أيضاً، وهذا بدوره يجعل تفسيرها من أجل بعض الحالات غير واضح أو غير ذي معنى، ولذلك اقترح تقديم مقياس آخر يتجاوز هذه السلبيات يدعى الدرجة المعياريّة -t (أو الدرجة المعيارية التائيّة)، ولكننا لن نتطرق إلى هذا المقياس في كتابنا هذا.



تمارين

١- عندما نتســوّق في أحد المتاجر للمواد التّموينيّة لاحظنا أنّه كُتِبَ على بعض الســلع عبارة "الوزن الصافي 2 كغ ± 100 غ"، فما هو تفسير هذه الكتابة؟

٢- إذا كانت x_1 و x_2 و... و x_n مع $1 \leq n$ مجموعة بيانات مُعطاة، فعندئذٍ:

أ- هل يمكن لقيمةٍ عدديةٍ x أن تكون مقياساً للنَّزعة المركزيَّة لهذه المجموعة من البيانات علماً أنَّها توافق x_ℓ (أكبر قيمة في البيانات) أو x_s (أصغر قيمة في البيانات) ؟

ب- بفرض أنَّ $\bar{x}$ هي قيمة المتوسط لهذه المجموعة من البيانات، فمن أجل أي قيمة لـ n تصبح $\bar{x} = x_\ell$ أو $\bar{x} = x_s$ ؟

ج- من أجل أيّة قيمة لـ n تصبح قيمة الانحراف المعيّاري لهذه المجموعة من البيانات تساوي الصفر؟

٣- لتكن لدينا مجموعة بيانات مرتَّبة لعينة مُعطاة كما يلي:

? 19 14 12 12 9 7 6 ? 3

فعندئذٍ:

أ- استخدم المقياس الملائم لتعيين القيمة التي تنزع إليها هذه البيانات.

ب- استخدم المقياس الملائم لتعيين القيمة التي تعبِّر عن تبعثر هذه البيانات حول مقياس نزعتها المركزية.

ج- هل يمكن لقيمة عددية x أن تكون مقياساً للنَّزعة المركزيَّة لهذه المجموعة من البيانات علماً أنَّها توافق x_ℓ (أكبر قيمة في البيانات) أو x_s (أصغر قيمة في البيانات) ؟

٤- لتكن لدينا مجموعة بيانات مرتَّبة لعينة مُعطاة كما يلي:

21 ? 17 15 12 ? 6 6 6 ? 3

فعندئذٍ:

أ- استخدم المقياس الملائم لتعيين القيمة التي تنزع إليها هذه البيانات.

ب- استخدم المقياس الملائم لتعيين القيمة التي تعبِّر عن تبعثر هذه البيانات حول مقياس نزعتها المركزية.

٥- لتكن لدينا مجموعة البيانات الآتية التي تمثِّل الوزن لعينة مكوَّنة من عشرين شخصاً بالغاً (مقدَّرةً بالكيلو غرام):

75	92	74	73	62	80	76	66	79	60
77	84	77	71	81	75	84	65	61	82

والمطلوب ما يلي:

أ- حساب: المنوال– الوسيط- المتوسط- الانحراف المعيّاري– مُعامل التغيُّر والدرجة المعيّارية للقيمة 75.

ب- تحت الفرض أنَّ لجميع قيم البيانات المُعطاة النصيب نفسه في الاختيار ولا يؤثِّر بعضها على البعض الآخر لدى عملية السحب، فعندئذٍ قم بسحب عينات عشوائية من البيانات بحجم $n = 5$ و $n = 10$ و $n = 15$، ومن ثمَّ احسب المتوسط والانحراف المعيّاري، وبعد ذلك قارن بين النتائج المتقابلة. ماذا تستنتج؟

٦- لتكن لدينا البيانات 8, 7, 6, 8, 4, 5, 9, 6, 9, 8 لعينة حجمها 10، والمطلوب حساب المدى الربيعي لهذه البيانات.

٧- ليكن لدينا مجموعة بيانات عينة مقدَّمة من خلال جدول توزيع تكراري له العرض الآتي:

التكرار المتجمِّع الصاعد للفئة	تكرار الفئة	مركز الفئة	حدود الفئة	رقم الفئة
	5		$50 \to 55$	1
13			$55 \to 60$	2
	12			3
40				4
	17			5
70				6
	7		$80 \to 85$	7
المجموع				**Total**

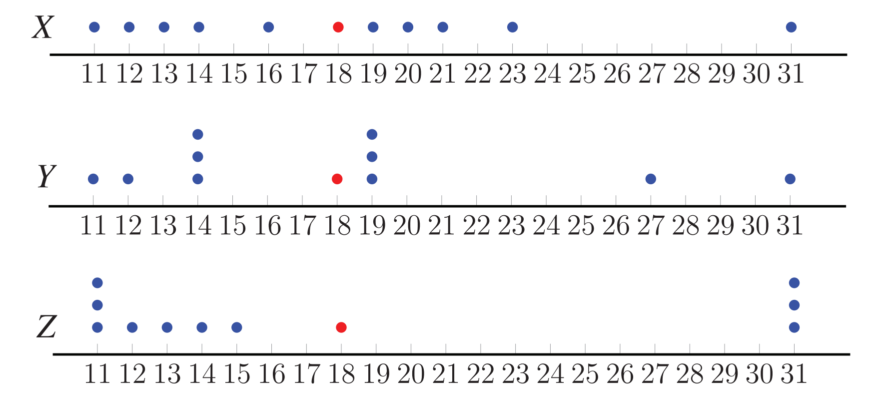
والمطلوب ما يلي:

أ- إكمال جدول التوزيع التكراري السابق.

ب- حساب كل من المتوسط، الوسيط والمنوال.

ج- حساب الانحراف المعيّاري.

٨- لتكن لدينا ثلاث مجموعات بيانات لعينات مقدَّمة كما في العروض الآتية:



والمطلوب ما يلي:

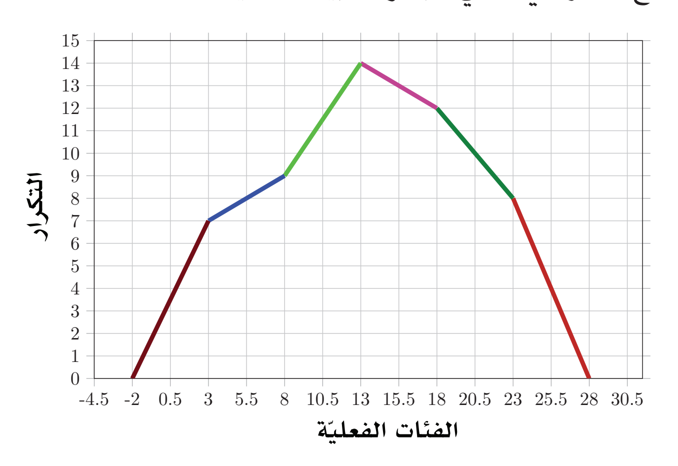
أ- حساب المتوسط، الوسيط، المدى، الانحراف المعيّاري والمدى الربيعي لكل من مجموعات البيانات المُعطاة.

ب- لو حُذِفَت القيمة $x = 18$ فهل تتغيَّر قيم المتوسطات لهذه المجموعات من البيانات، ولماذا؟

ج- حساب معامل التغيرُّ لكل من مجموعات البيانات المُعطاة، ومن ثمَّ بيِّن أيها أقل تبعثراً من الأخرى.

د- حساب الدرجة المعيّارية للقيمة $x = 14$ في كل من مجموعات البيانات المُعطاة، ماذا تلاحظ؟

٩- ليكن لدينا المضلَّع التكراري الآتي لمجموعة بيانات عينة:



والمطلوب ما يلي:

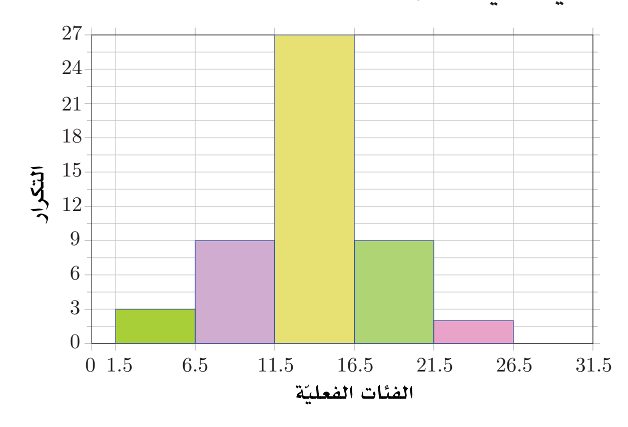
أ- بناء جدول التوزيع التكراري لبيانات هذا المضلَّع التكراري.

ب- حساب المتوسط، الوسيط والمنوال لبيانات هذا المضلَّع التكراري.

ج- حساب المدى لبيانات هذا المضلَّع التكراري.

د- حساب الانحراف المعيّاري لبيانات هذا المضلَّع التكراري.

٩- ليكن لدينا المدرَّج التكراري الآتي لمجموعة بيانات عيّنة:



والمطلوب ما يلي:

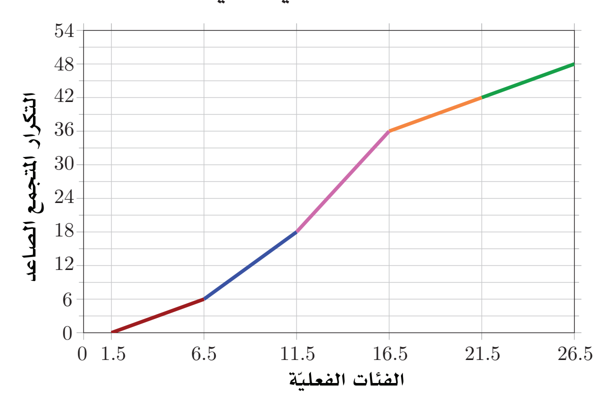
أ- بناء جدول التوزيع التكراري لبيانات هذا المدرَّج التكراري.

ب- حساب المتوسط، الوسيط والمنوال لبيانات هذا المدرّج التكراري.

ج- حساب المدى لبيانات هذا المدرّج التكراري.

د- حساب الانحراف المعيّاري لبيانات هذا المدرّج التكراري.

١٠- لتكن لدينا بيانات مقدّمة من خلال العرض البياني الآتي:



والمطلوب ما يلي:

أ- بناء جدول التوزيع التكراري للبيانات الخاصة بمضلَّع التكرار المتجمّع الصّاعد المعطى.

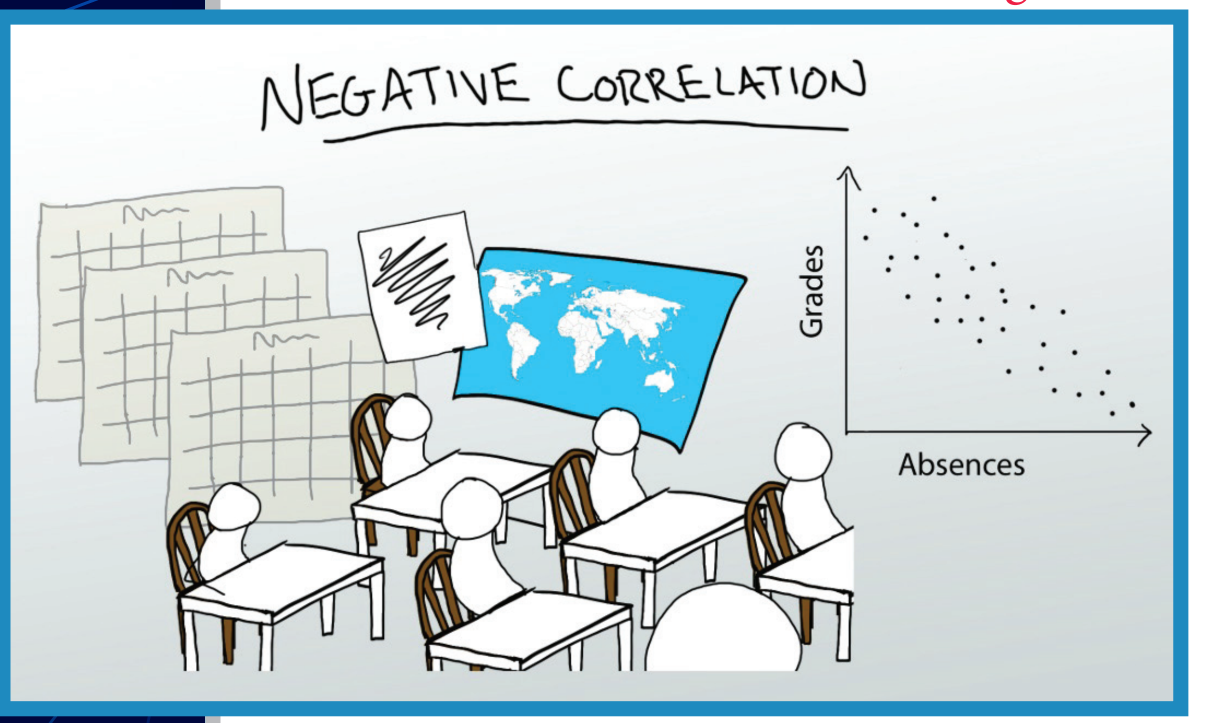
ب- حساب المتوسط، الوسيط والمنوال للبيانات الخاصة بمضلَّع التكرار المتجمّع الصّاعد المعطى.

ج- حساب المدى للبيانات الخاصة بمضلَّع التكرار المتجمّع الصّاعد المعطى.

د- حساب الانحراف المعيّاري للبيانات الخاصة بمضلَّع التكرار المتجمّع الصّاعد المعطى.

الفصل الرابع

الارتباط والانحدار الخطي
Linear Correlation and Regression



المقدمة:

لقد تحدثنا في الفصول الثلاثة السابقة عن طرائق إحصائية مُستخدمة في وصف ودراسة متغيرٍ واحد، ولاحظنا أنَّه يمكن لتلك الطرائق أن تُعطي تصوراً مقبولاً حول المتغيرِّ قيد الدراسة، ولكنَّها لا تقدِّم لنا تصوراً واضحاً ودقيقاً حول التّأثير المتبادل بين متغيّرين أو أكثر. فعلى سبيل المثال قد يتساءل المرء عن العلاقة التي تربط بين الطول والوزن لأشخاص من فئةٍ عمريةٍ معيّنة (فيكون لدينا متغيّرين)، أو بين تركيز النيكوتين في الدم وأمراض الرّئة وتصلّب الشرايين لدى مجتمع من المدخنين (فيكون لدينا ثلاثة متغيّرات)، أو مستوى الدخل واستهلاك بعض المواد الغذائية وزيادة الوزن واقتناء بعض أجهزة الرّفاهيّة (فيكون لدينا أربعة متغيّرات)، وهكذا دواليك .... فمن الممكن أن يتساءل المرء عن إمكانيّة تقدير أحد المتغيّرات إذا عُلِمتْ قيمة لمتغيّرٍ آخر. إنَّ العلم الذي يبحث في الإجابة على مثل هذه التساؤلات يُعرف باسم «تحليل الارتباط والانحدار» والذي سنقدِّم بعض مبادئه في هذا الفصل.

٤ - ١
الارتباط الخطي البسيط

قبل البدئ في دراستنا لهذا الفصل ننوّه إلى أنَّ دراستنا في هذا الفصل ستكون من أجل العيِّنات فقط، وكذلك سنتَّفق على استخدام كلمة متغيرٍّ للدلالة على ظاهرةٍ يُمثِّلها هذا المتغيرِّ، والذي بدوره يقوم بتوليد البيانات الخاصّة بهذه الظاهرة، فعلى سبيل المثال لو أخذنا الظاهرة A هي ظاهرة الطول لأشخاص عددهم n، فعندئذ المتغيرِّ X سيلعب دور وسيلة القياس (المسطرة) التي تقيس الطول وتولِّد البيانات x_1 و x_2 و... و x_n الممثَّلة لقيم الطول لهؤلاء الأشخاص، ويُقال في هذه الحالة إنَّ المتغيرِّ X راصدٌ للظاهرة A وواصفٌ للمجتمع الإحصائي الذي تنتمي إليه هذه البيانات.

الآن لنطرح السؤال الآتي:

هل التأثير المتبادل بين ظاهرتين مُمثَّلتين بمتغيرِّين X وY طردي أم عكسي، قوي أم ضعيف؟

من أجل الإجابة على هذه الأسئلة لا بدّ لنا من منهجٍ علميٍّ دقيقٍ وواضحٍ بعيداً عن التخمين الحدسي، ويقدّم لنا مقداراً عددياً يُعبرِّ عن طبيعة التّأثير المتبادل بين المتغيرّات. إنَّ الجانب الرّياضياتي الذي يهتّم بالإجابة على مثل هذه الأسئلة يُعرف باسم "**تحليل الارتباط**"، ويبحث هذا العلم في الكشف عن العلاقة بين متغيرٍّ واحد Y ومتغيرّات أخرى X_1 و X_2 و... و X_k، وفي هذه الحالة يُدعى Y **متغيرِّ الاستجابة** Response Variable وأمّا بقيّة المتغيرّات X_1 و X_2 و... و X_k فإنَّها تُدعى **متغيرّات تفسيرية** Explanatory Variables. في دراستنا هنا سنهتّم بشكل أساسيّ بالعلاقة التي تربط بين **متغيرِّ الاستجابة** Y و**متغيرِّ تفسيري** وحيد X، ويُدعى هذا النوع من الارتباط بـ **الارتباط البسيط** Simple Correlation.

الآن لنفترض أنَّه لدينا دراسةٌ متعلِّقةٌ برصد متغيرِّين (أو ظاهرتين) X وY فقـط، فعندئذ ستكون البيانات على شكل ثنائيات مرتبة $(x_i\,,\,y_i)$ من أجل كل القيم المكنة لـ i، ويُقال في هذه الحالة عن البيانات x_i و y_i إنَّها **بيانات مُتزاوجة** من أجل كل القيم المكنة لـ i، وهنا يُبْحث في اتجاهين رئيسيين أيضاً، وهما:

١- إذا كانت العلاقة التي تربط بين البيانات المتزاوجة x_i وy_i هي **علاقة خطيّة** من أجل كل القيم المكنة لـ i (**وفقاً لمفهوم الدالَّة الخطيّة التي قدّمت في الفصل التمهيدي**).

٢- إذا كانت العلاقة التي تربط بين البيانات المتزاوجة x_i و y_i هي **علاقة غير خطية** من أجل كل القيم المكنة لـ i (**كأن تكون العلاقة بينهما على شكل دالَّة من الدرجة الثانية وما فوق مثلاً**).

٤-١-١- بعض نماذج الارتباط

فيما يلي سنهتّم بدراسة العلاقة بين متغيرِّين (أو ظاهرتين) X وY تربط بينها **علاقة خطية** فقـط، وكتمهيد لذلك سنقدِّم بعض نماذج الارتباط بين المتغيرّات التي أساسها **التأثير والتأثُّر**، حيث يوجد في هذا المضمار عدَّة أنواع من الارتباط بحسب العلاقة التي تصفه، ومن هذه النماذج ما يلي:

أ- الارتباط السببي Causal Correlation:

هذا النوع من الارتباط مبني على أنَّ التغيُّر في قيم متغيِّر يؤدي إلى تغيُّر في قيم متغيِّر آخر مرتبط به، وهذا التأثير غير عكوس. أي أنَّ التغيُّر في المتغيِّر الأخير لا يؤثِّر في تغيُّر المتغيِّر الأول، ولذلك يُدعى هذا النوع من الارتباط بـ **الارتباط السببي**، ومن الأمثلة على ذلك:

١- زيادة نسبة الأسمدة وفق معايير سليمة يؤدي إلى زيادة الإنتاج، إلاّ أنَّ العكس ليس صحيحاً بالضرورة.

٢- كلما ازداد ارتفاعنا عن مستوى سطح البحر فإنَّ درجة الحرارة ستنخفض عموما، إلاّ أنَّ العكس ليس صحيحاً بالضرورة.

ب- ارتباط السلسلة السببية Causal Chain Correlation:

إنَّ هذا النموذج من الارتباط يشابه النموذج السابق من الارتباط، ولكنَّه يكون في سلسلة من نماذج الارتباط السببي، فعلى سبيل المثال:

هطول الأمطار يؤدي إلى تسرب مياه الأمطار إلى التّربة، وهذا يزيد المحتوى المائي للتربة، والتي بدورها تتسرب إلى مخازن المياه الجوفية، ونحصل نتيجة لذلك على زيادة في مخزون المياه الجوفية. إنَّ هذا النوع من الارتباط يُمكن إدراجه تحت اسم الارتباط السببي أيضاً.

ج- الارتباط التبادلي Reciprocal Correlation:

إنَّ هذا النوع من الارتباط مبني على التأثير المتبادل بين المتغيِّرات أو على ردود الأفعال بين المتغيِّرات، فكل تغيُّر في متغيِّر يؤدي إلى تغيُّر في المتغيِّر الآخر، فعلى سبيل المثال لو أخذنا حادثة تصادم جسمين يسيران بسرعة ما (**غير ساكنين**) فإنَّ كل جسم سيوثِّر في الجسم الآخر بقوة تتناسب مع كتلته وسرعته عند لحظة التصادم. إنَّ هذا النوع من الارتباط يُعرف بـ **الارتباط التبادلي** Reciprocal Correlation.

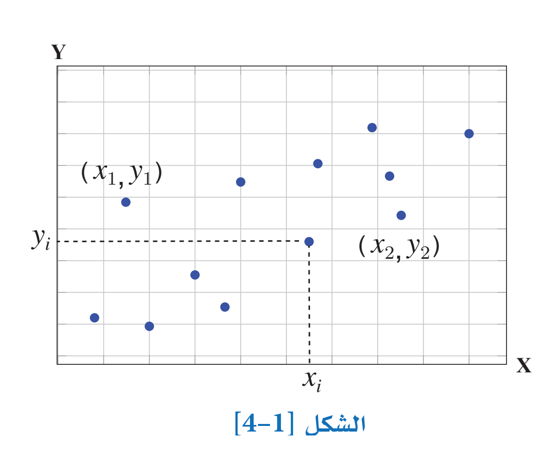
د- الارتباط الوهمي Spurious Correlation:

في هذا النوع من الارتباط يبدو لنا ظاهرياً نوع من التأثير لمتغيِّر على متغيِّر آخر، ولكنَّ هذا التأثير غير حقيقي، فعلى سبيل المثال لو أخذنا تغيُّر مستوى وعي الطفل ومقارنته مع تغيُّر طول ذراعه، فتجد أنَّ هناك علاقة طردية بينهما ولكنَّها ليست حقيقية، فلا تأثير لطول ذراع الطفل على وعيه ولا لوعيه على طول ذراعه، ولذلك يُعرف هذا النوع من الارتباط بـ **الارتباط الوهمي** Spurious Correlation.

الآن، وبعد هذا التقديم ننتقل إلى النمذجة الرياضياتية للارتباط بين متغيِّرين X و Y، والمتمثِّلة بتقديم بعض المعايير العددية التي تقيس لنا درجة الارتباط بين هذين المتغيِّرين. إنَّ دراسة الارتباط بين متغيِّرين X و Y تبدأ عادة بما يُسمّى **لوحة الانتشار**، حيث تعدّ لوحة الانتشار من الوسائط الهامَّة في دراسات الارتباط. إنَّ الغاية من هذه اللوحة أخذ انطباع أولي عن طبيعة التأثير المتبادل بين المتغيِّر التفسيري X ومتغيِّر الاستجابة Y.

٤-١-٢- لوحة الانتشار Scatter Plot

لتكن لدينا (x_1, y_1) و (x_2, y_2) و..... و (x_n, y_n) بيانات مُعطاة، فعندئذ لوحة الانتشار لهذه البيانات هي تمثيل نقطي للبيانات (أي تمثيل البيانات بنقاط) على المستوي الإحداثي XoY، ويُؤخذ عادةً في الإحداثيات المتعامدة، أي أنَّ x_1 و x_2 و... و x_n و y_1 و y_2 و... و y_n هي مساقط النقاط الممثّلة لهذه البيانات على المحور oX و oY على الترتيب (انظر الشكل [4–1])، ولإعطاء المزيد من التوضيح سنقدِّم الأمثلة الآتية.



الشكل [4–1]

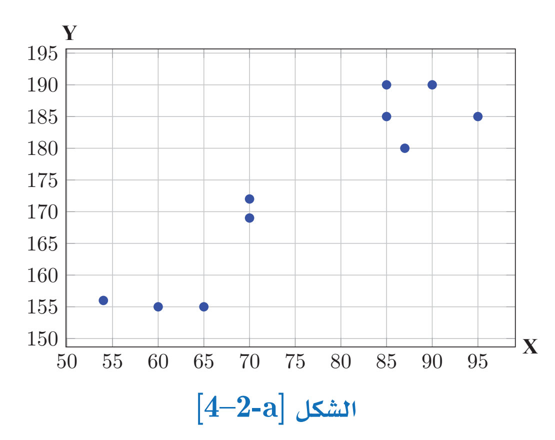
٤-١-٢-١- أمثلة

١- قمنا بقياس الطول (مقدّراً بالسنتيمتر) والوزن (مقدّراً بالكيلوغرام) لعشرة أشخاص بالغين فحصلنا على البيانات المدوّنة في الجدول الآتي:

الجدول [4–1]

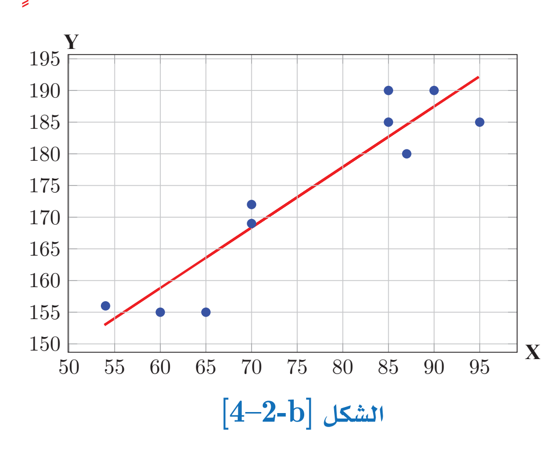
الطول Y	الوزن X	i	الطول Y	الوزن X	i
169	70	6	155	60	1
185	85	7	156	54	2
155	65	8	190	85	3
190	90	9	185	95	4
180	87	10	172	70	5

عندئذ نجد أنَّ لوحة الانتشار لهذه البيانات لها العرض الموضَّح بالشكل الآتي.



الشكل [4–2-a]

فنلاحظ أنَّ ثمّة علاقة تربط بين هاتين المجموعتين من البيانات، حيث يبدو لنا جلياً أنَّ النقاط الممثّلة للبيانات (x_1, y_1) و (x_2, y_2) و... و (x_{10}, y_{10}) تُظْهِر مَنْحىً صاعد، بمعنى أنَّه كلما ازدادت قيمة مُركّبة أدى ذلك إلى زيادةٍ في قيمة المُركّبة الأخرى، وهذا السلوك في الارتباط يُدعى **الارتباط الإيجابي** (أو الطردي) للبيانات. أمّا بخصوص قوَّة التأثير المتبادل (أي قوّة الارتباط، وسوف نوضِّحه لاحقاً) بين الطول والوزن، فإنَّ الشكل يظهر لنا درجةً عالية من التأثير المتبادل بين الطول والوزن بسبب وقوع معظم البيانات بالقرب من مستقيم مُمثِّل لمنحى هذه البيانات (سنأتي على شرح كيفية تعيين هذا المستقيم لاحقاً في بحث الانحدار الخطي من هذا الفصل)، انظر الشكل المجاور [4-2-b] .



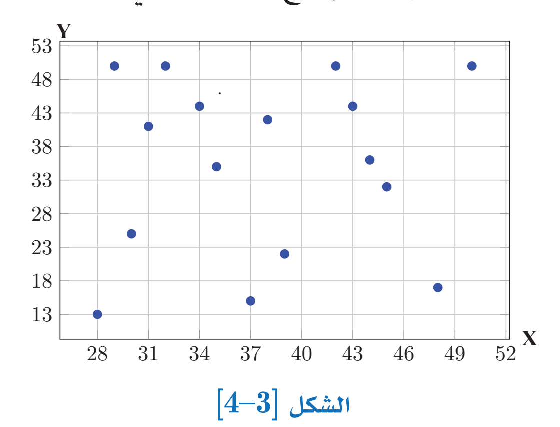
الشكل [4–2-b]

٢- تقدَّم ستة عشر طالباً للاختبار النهائي في مقرَّري الرياضيات واللغة العربية، فكانت لهم النتائج الآتية مقدَّرةً من 50 درجة.

الجدول [4–2]

i	درجة اختبار الرياضيات X	درجة اختبار اللغة العربية Y	i	درجة اختبار الرياضيات X	درجة اختبار اللغة العربية Y
1	45	32	9	28	13
2	48	17	10	35	35
3	30	25	11	39	22
4	50	50	12	42	50
5	43	44	13	32	50
6	37	15	14	38	42
7	34	44	15	31	41
8	44	36	16	29	50

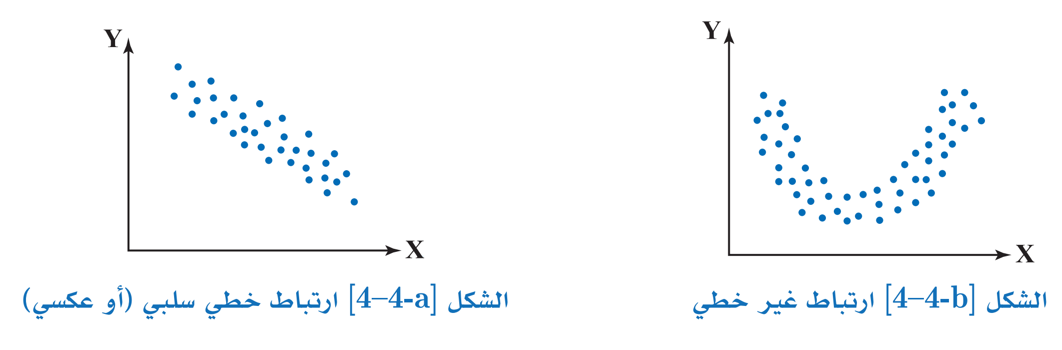
فنجد لوحة الانتشار لهذه البيانات لها العرض الموضَّح بالشكل الآتي.



الشكل [4–3]

وهنا نلاحظ أنَّ النقاط الممثّلة للبيانات المُعطاة تُظهر توزعاً على شكل غيمةٍ يصعب تحديد اتجاهها (**المَنْحى غير واضح تماماً**)، بمعنى أنَّه من خلال النظر بالعين المجردة يصعب تحديد طبيعة التأثير المتبادل بين المتغيّرين X وY من حيث الإيجابيّة أو السلبيّة.

من النماذج الأخرى التي تصف العلاقة بين متغيّرين يوجد الارتباط السلبيّ (**أو العكسيّ**) وكذلك الارتباط غير الخطيّ، والشكلين الآتيين يوضّحان ذلك.



الشكل [4−4-b] ارتباط غير خطي

الشكل [4−4-a] ارتباط خطي سلبي (أو عكسي)

الآن، وبما أنَّ النظر إلى الأشكال البيانية يختلف من شخص إلى آخر فإنَّه لا بدَّ من البحث عن مقياس عددي يوضّح لنا طبيعة الارتباط أهو إيجابي أم سلبي، وإن كان قوياً أم ضعيفاً. إنَّ المقياس العددي الذي يقوم بهذه المهمَّة يُدعى **مُعامل الارتباط** Correlation Coefficient، وسنرمز له بـ **r**. لكن قبل القيام بتقديم بعض المفاهيم المتعلّقة بالارتباط والانحدار، فإنَّنا سننوّه إلى الآتي تجنباً للتقديم المطوَّل، فعندما نذكر مستقبلاً أنَّه لدينا (x_1,y_1) و (x_2,y_2) و… و (x_n,y_n) بيانات عيّنة خاضعة لمتغيّرين X وY، فإنَّنا نقصد بذلك:

أ- إمّا أن تكون عناصر عيّنة حجمها n مسحوبة من مجتمع إحصائي أخضعت لدراسة هذين المتغيّرين X وY وحصلنا نتيجةً لذلك على البيانات المتزاوجة المذكورة أعلاه.

ب- أو أن تكون البيانات x_1 و x_2 و… و x_n هي قيم ناتجة عن إخضاع عناصر عيّنة حجمها n مسحوبة من مجتمع إحصائي لدراسة المتغيّر X، وكذلك البيانات y_1 و y_2 و… و y_n هي قيم ناتجة عن إخضاع عناصر عيّنة حجمها n مسحوبة من مجتمع إحصائي آخر لدراسة المتغيّر Y، وبحيث تُكوِّن x_1 و x_2 و… و x_n وكذلك y_1 و y_2 و… و y_n البيانات المتزاوجة المذكورة أعلاه.

٤-١-٣- مُعامل بيرسون للارتباط Person's Coefficient of Correlation

لتكن (x_1,y_1) و (x_2,y_2) و… و (x_n,y_n) بيانات عيّنة أخضعت لمتغيّرين X و Y، وأنَّ للبيانات x_1 و x_2 و… و x_n متوسط وانحراف معياري $\bar{x}$ و S_X على الترتيب، وكذلك للبيانات y_1 و y_2 و… و y_n متوسط وانحراف معياري $\bar{y}$ و S_Y على الترتيب، فإذا كان $0 < S_X$ وكذلك $0 < S_Y$ فعندئذ يُعطى مُعامل الارتباط الخطي لهذه البيانات بالعلاقة الآتية:

$$\mathbf{r} = \frac{1}{n-1}\sum_{i=1}^{n}\frac{(x_i-\bar{x})}{S_X}\cdot\frac{(y_i-\bar{y})}{S_Y}$$ [4–1-a]

إنَّ هذه العلاقة تُعرف باسم "**مُعامل بيرسون للارتباط الخطي**" نسبةً إلى الإحصائي الإنجليزية بيرسون Karl Pearson (1875-1936) والذي كان رائداً في الإحصاء، وكما هو ملاحظ يمكن استخدام هذه العلاقة إذا كانت قيم المتوسطات والانحرافات المعيَّارية للبيانات مُعطاة. أمَّا إذا كانت قيم الانحرافات المعيَّارية للبيانات ليست مُعطاة ولدينا قيم المتوسطات للبيانات مُعطاة فقط، فعندئذ يمكننا استخدام العلاقة الآتية التي تنتج عن العلاقة السابقة باستخدام الصيغة التي تعطي قيمة الانحراف المعيَّاري بدلالة قيم البيانات والمتوسط (انظر العلاقة [3–1-a] في الفصل السابق):

$$\mathbf{r} = \frac{\sum_{i=1}^{n} x_i\, y_i - n\,\bar{x}\,\bar{y}}{\sqrt{\sum_{i=1}^{n} x_i^2 - n\,\bar{x}^2}\cdot\sqrt{\sum_{i=1}^{n} y_i^2 - n\,\bar{y}^2}}$$ [4–1-b]

وأخيراً إذا كانت قيم المتوسطات ليست مُعطاة فإنَّه يمكننا استخدام قيم البيانات مباشرة في حساب قيمة مُعامل الارتباط من خلال العلاقة الآتية التي تنتج عن العلاقة السابقة بعد التعويض عن المتوسطات بما يساويها بدلالة مجاميع البيانات.

$$\mathbf{r} = \frac{n\cdot\left(\sum_{i=1}^{n} x_i y_i\right)-\left(\sum_{i=1}^{n} x_i\right)\cdot\left(\sum_{i=1}^{n} y_i\right)}{\sqrt{n\cdot\left(\sum_{i=1}^{n} x_i^2\right)-\left(\sum_{i=1}^{n} x_i\right)^2}\cdot\sqrt{n\cdot\left(\sum_{i=1}^{n} y_i^2\right)-\left(\sum_{i=1}^{n} y_i\right)^2}}$$ [4–1-c]

٤-١-٣-١- ملاحظات

١- إذا كان $S_X = 0$ أو $S_Y = 0$ فعندئذ نضع بالتعريف $\mathbf{r} = 0$.

٢- يُستخدم هذا المقياس من أجل البيانات الخام الكميّة فقط، ويُعدّ من المقاييس الجيدة للارتباط الخطي، وجودته تكمن في أنَّه يستخدم جميع قيم بيانات المتغيرِّين X وY في الحساب.

٣- إنَّ مُربَّع قيمة مُعامل الارتباط تُدعى **مُعامل التحديد** Coefficient of Determination، ويُرمز له بـ $\mathbf{r}^2$ (وفي بعض البرامج الإحصائية يُشار إليه بـ **R-square**)، وهذه القيمة تستعمل كتقدير لقوة علاقة الارتباط بين متغيرِّين أيضاً.

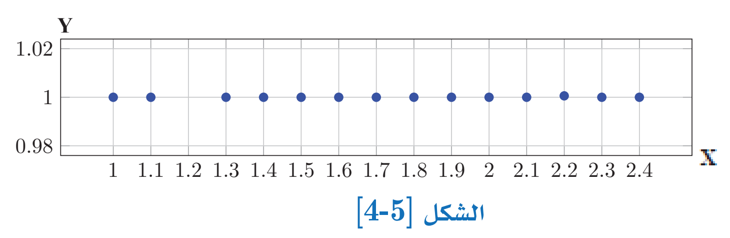
٤-١-٤- خصائص مُعامل الارتباط الخطي

إنَّ لمعامل الارتباط الخطي **r** خصائص تميّزه، ومن أهمّها ما يلي:

١- إنَّ قيمة **r** تنتمي للفترة $[-1 , +1]$ دوماً، أي أنَّ $-1 \le \mathbf{r} \le +1$.

٢- إذا كان $\mathbf{r} > 0$ فعندئذ يُقال إنَّ الارتباط بين المتغيِّرين X وY إيجابي (أو طردي)، وأمّا إذا كان $\mathbf{r} < 0$ فإنَّه يُقال إنَّ الارتباط بين المتغيِّرين X وY سلبي (أو عكسي).

٣- إذا كان $\mathbf{r} = \pm 1$ فعندئذ يُقال عن الارتباط بين المتغيِّرين X وY إنّه خطي تام، وفي هذه الحالة تكون جميع النقاط المُمثِّلة للمشاهدات (x_i , y_i) من أجل كل القيم المكنة لـ i تقع على خطٍ مستقيمٍ واحد. لكن يجب التنبيه هنا إلى أنَّ العكس ليس صحيحاً بالضرورة، فلو نظرنا إلى مجموعة البيانات الموضَّحة من خلال الشكل الآتي لوجدنا أنَّ جميع النقاط المُمثِّلة لها تقع على مستقيمٍ واحدٍ، ولكن بسبب انعدام الانحراف المعيَّاري S_Y (أي $S_Y = 0$) فإنَّه ستكون قيمة مُعامل الارتباط الخطي $\mathbf{r} = 0$.



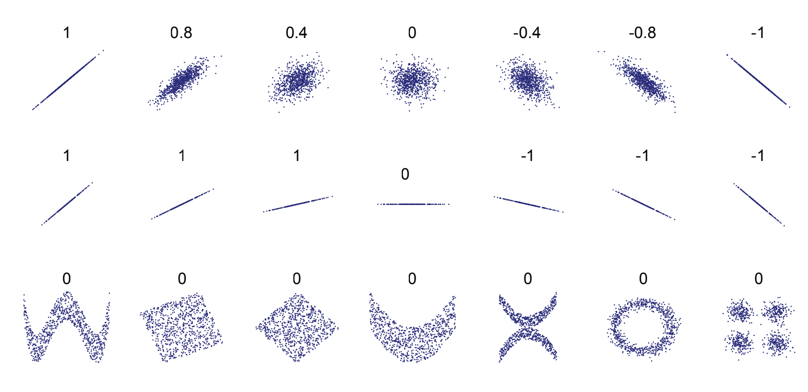
الشكل [4-5]

٤- إذا كان $\mathbf{r} = 0$ ، فعندئذ **لا يُقال** إنَّه لا يوجد ارتباط، بل نكون أمام إحدى الحالات الثلاث الآتية:

أ- إمّا أن يكون توزّع البيانات على لوحة الانتشار على شكل غيمة عديمة الاتجاه، أو في تجمعات جزئية متناثرة تشكِّل بمجملها مشهداً يصعب تحديد اتجاهه، ومن ثمَّ مَنْحى البيانات سيكون غير واضحٍ، ومن الأمثلة على ذلك تجدها في الشكل الآتي [4-6].

ب- أو أن يكون مَنْحى البيانات موازياً لأحد محوري لوحة الانتشار، ومن الأمثلة على ذلك الشكل [4-5]، وتجد بعض العروض التوضيحية الأخرى في الشكل الآتي [4-6].

ج- وأخيراً من الممكن أن يكون توزّع البيانات على لوحة الانتشار له شكل بيان علاقةٍ غير خطيّة بين المتغيِّرين X وY ، ونماذج توضيحية على ذلك تجدها في الشكل الآتي [4-6] أيضاً.



الشكل [4-6]

٤-١-٥- تقييم قوَّة وضعف الارتباط الخطيّ

لتقييم قوّة وضعف الارتباط الخطي بين بيانات متغيّرين X وY، فإنَّه يُقال عن الارتباط الخطي بينهما إنَّه:

أ- قوي جداً إذا كان $0.86 < |\mathbf{r}| \leq 1$، وإذا كان $\mathbf{r} = \pm 1$ فعندئذ يكون الارتباط بين بيانات المتغيِّرين X وY خطي تام، وقوة الارتباط تكون في أوجْها (قد بلغت قيمتها العظمى).

ب- قوي إذا كان $0.70 < |\mathbf{r}| \leq 0.86$،

ج- متوسط القوّة إذا كان $0.50 < |\mathbf{r}| \leq 0.70$،

د- ضعيف إذا كان $0.30 < |\mathbf{r}| \leq 0.50$،

هـ- ضعيف جداً إذا كان $0 \leq |\mathbf{r}| \leq 0.30$.

٤-١-٦- أمثلة

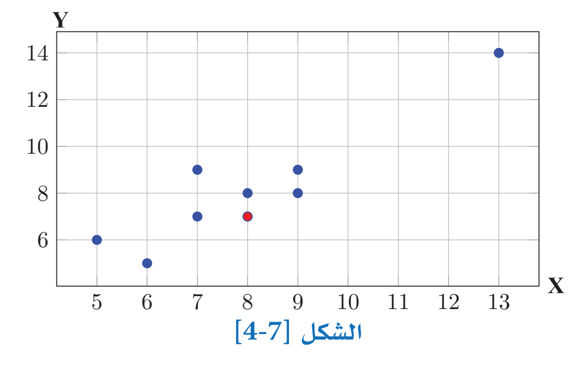
تمَّ تدريب عشرة طلاب على نوع من التمارين الرياضية، بحيث يُدرَّب الطالب ذي الرقم k لعدد من المرّات يساوي x_k مرَّة، ولنفترض أنَّ X متغيِّر يرصد التكرارات التي قام بها الطلاب. بعد ذلك يَخضع كل طالب لاختبار فيمكن له أن يحصل على درجة y_k من أصل 15 درجة (أي أنَّ الدرجة القصوى 15)، ولنفترض أنَّ Y متغيِّر يرصد تقييم الطالب. الآن بفرض أنَّ نتائج الطلاب كانت كما في الجدول الآتي:

الجدول [4-3-a]

رقم الطالب i	10	9	8	7	6	5	4	3	2	1
عدد مرّات تدريب الطالب X	8	6	8	13	8	7	9	9	5	7
درجة اختبار الطالب Y	8	5	7	14	7	7	8	9	6	9

فهل تشير هذه النتائج إلى وجود ارتباط خطي بين المتغيِّرين X وY (أي بين عدد مرّات التدريب والنتيجة التي حصل عليها الطالب)، وما قيمته؟

لنقم أولاً بعرض لوحة الانتشار للبيانات فنجد لها العرض الجانبي، ونشير هنا إلى أنَّنا قمنا بتمييز النقطة المضاعفة $(8,7)$ بنقطة زرقاء وفي داخلها نقطة حمراء.



الشكل [4-7]

فنلاحظ وجود ارتباط إيجابي بين عدد مرَّات التدريب ودرجة الاختبار، وهذا يعني أنَّه كلما تلقى الطالب دورات تدريبية أكثر كلما حصل على نتائج أفضل، وأمّا من أجل حساب معامل بيرسون للارتباط الخطي فإنَّنا سنقوم بتقديم الجدول الآتي الذي يسهِّل علينا عملية الحساب لهذا المُعامل.

الجدول [4-3-b]

i	x_i	x_i^2	y_i	y_i^2	$x_i \cdot y_i$
1	7	49	9	81	63
2	5	25	6	36	30
3	9	81	9	81	81
4	9	81	8	64	72
5	7	49	7	49	49
6	8	64	7	49	56
7	13	169	14	196	182
8	8	64	7	49	56
9	6	36	5	25	30
10	8	64	8	64	64
المجموع	**80**	**682**	**80**	**694**	**683**

فنجد أنَّ قيمة معامل بيرسون للارتباط الخطي بين بيانات X وY تساوي:

$$\mathbf{r} = \frac{n \cdot \left(\sum_{i=1}^{n} x_i y_i\right) - \left(\sum_{i=1}^{n} x_i\right) \cdot \left(\sum_{i=1}^{n} y_i\right)}{\sqrt{n \cdot \left(\sum_{i=1}^{n} x_i^2\right) - \left(\sum_{i=1}^{n} x_i\right)^2} \cdot \sqrt{n \cdot \left(\sum_{i=1}^{n} y_i^2\right) - \left(\sum_{i=1}^{n} y_i\right)^2}}$$

$$= \frac{10(683) - (80 \times 80)}{\sqrt{10(682) - (80)^2}\ \sqrt{10(694) - (80)^2}} = \frac{430}{476.24} = 0.903$$

وهذه النتيجة تؤكّد لنا مرّة أخرى على وجود ارتباط **إيجابي قوي جداً** بين عدد مرّات التدريب ودرجة الاختبار.

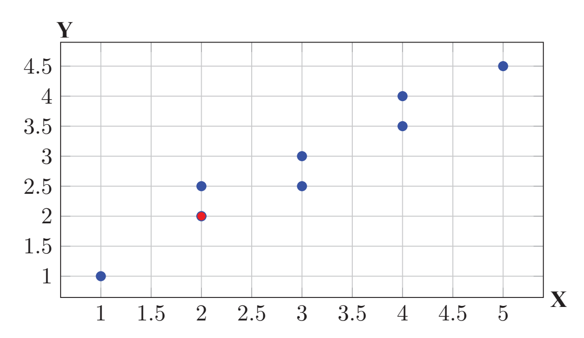
٢- الجدول الآتي يعرض لنا عدد الساعات الدراسيّة التي يمضيها الطالب في دراسة مقرّراته (**استذكار الطالب لمقرّراته**) في السنة الأولى المشتركة وما يقابلها من درجة التحصيل الفصلي، علماً أنَّ X متغيِّر يرصد عدد الساعات الدراسيّة في حين Y متغيِّر يرصد درجة التحصيل الفصلي للطالب.

الجدول [4-4-a]

رقم الطالب i	9	8	7	6	5	4	3	2	1
عدد ساعات الدراسة X	4	3	2	5	4	2	1	3	2
درجة التحصيل الفصلي Y	3.5	3	2	4.5	4	2.5	1	2.5	2

فهل تشير هذه النتائج إلى وجود ارتباط خطي بين المتغيِّرين X وY (أي بين عدد ساعات الدراسة ودرجة التحصيل الفصلي للطالب)، وما قيمته؟

لنقم أولاً بعرض لوحة الانتشار للبيانات فنجد لها العرض الآتي، ونشير هنا إلى أنَّنا قمنا بتمييز النقطة المضاعفة $(2,2)$ بنقطة زرقاء وفي داخلها نقطة حمراء.



الشكل [4-8]: العرض الانتشاري للبيانات المقدَّمة في الجدول [4-4-a]

فنلاحظ وجود ارتباط إيجابي بين عدد ساعات الدراسة ودرجة التحصيل الفصلي، وهذا يعني أنَّه كلما ذاكر الطالب لعدد ساعات أكثر كلما حصل على درجة أعلى في نتائج التحصيل الفصلي، وأمّا من أجل حساب معامل بيرسون للارتباط الخطي فإنَّنا سنقوم بتقديم الجدول الآتي الذي يسهِّل علينا عملية الحساب لهذا المُعامل.

الجدول [4-4-b]

i	x_i	x_i^2	y_i	y_i^2	$x_i \cdot y_i$
1	2	4	2	4	4
2	3	9	2.5	6.25	7.5
3	1	1	1	1	1
4	2	4	2.5	6.25	5
5	4	16	4	16	16
6	5	25	4.5	20.25	22.5
7	2	4	2	4	4
8	3	9	3	9	9
9	4	16	3.5	12.25	14
المجموع	**26**	**88**	**25**	**79**	**83**

فنجد أنَّ قيمة معامل بيرسون للارتباط الخطي بين بيانات X وY تساوي:

$$\mathbf{r} = \frac{n \cdot \left(\sum_{i=1}^{n} x_i y_i\right) - \left(\sum_{i=1}^{n} x_i\right) \cdot \left(\sum_{i=1}^{n} y_i\right)}{\sqrt{n \cdot \left(\sum_{i=1}^{n} x_i^2\right) - \left(\sum_{i=1}^{n} x_i\right)^2} \cdot \sqrt{n \cdot \left(\sum_{i=1}^{n} y_i^2\right) - \left(\sum_{i=1}^{n} y_i\right)^2}}$$

$$= \frac{9(83) - (26 \times 25)}{\sqrt{9(88) - (26)^2}\ \sqrt{9(79) - (25)^2}} = \frac{97}{99.84} = 0.97$$

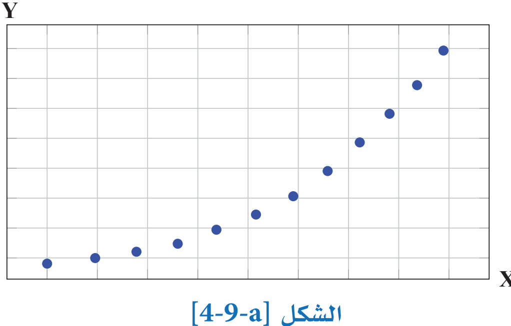
وهذه النتيجة تؤكِّد لنا ثانية وجود علاقة ارتباط خطيّة **إيجابية قويّة جداً** بين عدد ساعات الدراسة الأسبوعيّة والمعدّل الفصلي الذي حصل عليه الطالب.

٤ - ٢
الانحدار الخطي البسيط

إنَّ أصل مفهوم الانحدار جاء (في أواخر القرن التاسع عشر) من دراسات في علم الوراثة للإحصائي وعالم الوراثة والنفس والاجتماع الإنجليزية غالتون (1822-1911) Sir Francis Galton عند تقديم دراسته حول العلاقة التي تربط بين أعمار الآباء والأبناء لأجيال متعاقبة عديدة، وذلك بغية استقراء واقع تطور العمر لدى العنصر البشري، فوجد أنَّ أعمار الأبناء في حالة انحسار متواصل عبر الأجيال المتعاقبة، وكذلك لاحظ غالتون أن خصائص محدَّدة مثل الطول في الآباء لا تنتقل تماماً إلى ذرياتهم (الطول في حالة انحسار أيضاً)، وعُبِّر عن ذلك بالقول: إنَّ أعمار الأبناء في حالة انحدار لأنَّ التمثيل البياني لها كان منحدراً، ومن هنا جاءت تسمية تحليل الانحدار (ومن ثمَّ منشأ التسميّة تاريخي ولا علاقة له بسلوك البيانات منحدرة كانت أم صاعدة).

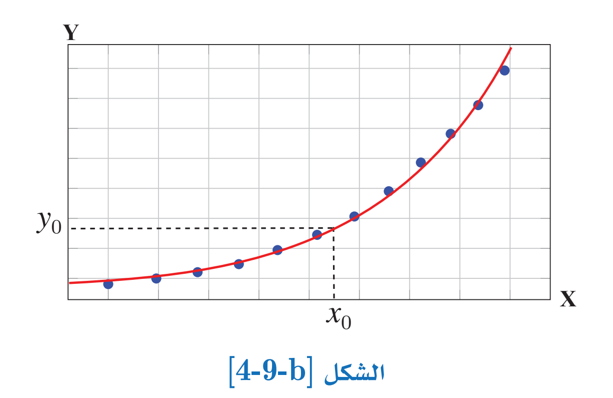
إنَّ تقدير قيمة ما لمتغيِّر X إذا ما عُلمت علاقة ارتباطه بمتغيِّر آخر Y مبنيّة على ما يُسمَّى بتحليل الانحدار. في الواقع إنَّ الانحدار هو تمثيل للعلاقة المتوسطة بين المتغيِّرات، فإذا وجدت علاقة تربط بين متغيِّرين مطلوب دراستهما فإنَّه يمكن توفيق معادلة لمنحنٍ (أو لخطٍ) يُحدِّد طبيعة تلك العلاقة، ومن ثمَّ يمكننا استخدام ذلك المنحني لتقدير القيم النظرية لمتغيِّر إذا عُلمت القيم للمتغيِّر الآخر.

الآن، ولتوضيح ذلك رياضياتياً سنفترض أنَّه لدينا (x_1, y_1) و (x_2, y_2) و ... و (x_n, y_n) مجموعة منتهية من المشاهدات الناتجة عن متغيِّرين X و Y، ولنفترض على سبيل المثال أنَّ تمثيلها على لوحة الانتشار له العرض المقدَّم في الشكل المجاور [4-9-a].



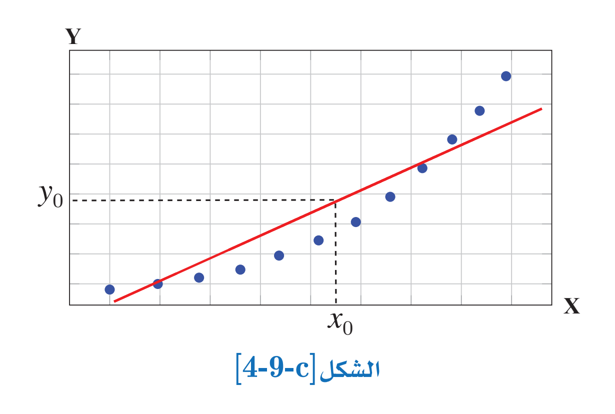
الشكل [4-9-a]

فعندئذ من أولى مهام تحليل الانحدار تمكيننا من تقدير قيمة y_o مقابلة لقيمة x_o مجهولة تقع ضمن نطاق البيانات الأصلية للمتغيِّر X. بالطبع إنَّ أفضل تقدير لهذه القيمة ينتج عن استخدام الدالَّة التي يمرّ رسمها البياني من جميع النقاط الممثَّلة للبيانات، ولكن تعيين معادلة مثل هذه الدالَّة في الحالة العامّة قد يكون صعب جداً إن لم يكن غير ممكن في بعض الحالات، ولذلك يبحث المرء في تعيين الدالَّة التي رسمها البياني يمرّ من بعض النقاط أو بالقرب من بعضها الآخر إن أمكن ذلك (انظر الشكل [4-9-b]).



الشكل [4-9-b]

ولكن تعيين الدوالّ غير الخطية قد يحتاج في كثير من الأحيان إلى برامج خاصّة بتوفيق **المنحنيات** Curves Fitting الأمر الذي قد لا يكون متوفراً إلاَّ لدى الاختصاصيين. لذلك سنبحث فيما يلي في تعيين معادلة أبسط أنواع المنحنيات (**ألا وهي المستقيمات**) التي تمرّ من بعض هذه النقاط وبالقرب من بعضها الآخر إن أمكن ذلك. إنَّ هذه المعادلة تُدعى **معادلة الانحدار الخطي البسيط**. أي أنَّنا سنسعى إلى تعيين المستقيم الذي يمرّ من جميع هذه النقاط أو يمرّ من بعض هذه النقاط وبالقرب من بعضها الآخر إن أمكن ذلك، وإن لم يكن ذلك ممكناً فليكن المستقيم الذي يمرّ بالقرب من هذه النقاط كما في الشكل الآتي [4-9-c]. إنَّ هذا المستقيم يُدعى **مستقيم الانحدار**.



الشكل [4-9-c]

في الواقع توجد طرائق عديدة لتعيين العـلاقة الخطيّة التي تربط بين متغيِّرين X و Y، ولكنَّنا سنقدِّم معادلة مستقيم الانحدار اعتماداً على ما يُسمَّى بـ **"طريقة المربّعات الصُغْرى"** (**لن نتناول شرحها**) فقط.

٤-٢-١- معادلة مستقيم انحدار Y على X

Equation of Straight Regression of Y on X

لتكن لدينا (x_1,y_1) و (x_2,y_2) و ... و (x_n,y_n) بيانات عيّنة، فإذا كنّا نرغب في تعيين قيمة $\hat{y}_i$ كتقدير لقيمة y_i باستخدام مستقيم الانحدار لدى إعطاء قيمة محدّدة لـ x_i، فإنَّه يمكننا ذلك باستخدام معادلة مستقيم الانحدار لهذه البيانات (**وفقاً لطريقة المربعات الصغرى**) التي لها العرض الآتي:

$$\hat{Y} = a + b\,X \qquad [4\text{–}2]$$

علماً أنَّ a وb ثوابت حقيقية يتمّ حسابهما كما يلي:

أ- باستخدام قيم البيانات مباشرة حيث نضع:

$$b = \frac{n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2} \qquad [4\text{–}3]$$

$$a = \frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2} \qquad [4\text{–}4\text{-a}]$$

ب- إذا كنّا نعلم قيم المتوسطين $\bar{x}$ و $\bar{y}$ فإنَّنا نحسب b من العلاقة [4–3]، وأمّا قيمة a فإنَّها تُحسب من العلاقة الآتية:

$$a = \bar{y} - b\bar{x} \qquad [4\text{–}4\text{-b}]$$

وهكذا تكون معادلة المستقيم [4–2] معيَّنة تماماً بعد حساب قيم الثابتين a وb.

إنَّ المستقيم الناتج عن المعادلة [4–2] يُدعى **مستقيم انحدار** Y على X للبيانات المُعطاة، ويُنظر إلى هذا المستقيم على أنَّه أفضل مستقيم يمرّ بالنقاط المثِّلة لهذه البيانات، وهذه العلاقة تساعدنا في تقدير (أو **التخمين عن**) القيمة $\hat{y}_o$ الموافقة لقيمة x_o تقع **ضمن** نطاق بيانات المتغيِّر X، وفي هذه الحالة تُدعى القيمة $\hat{y}_o$ بالقيمة المُقدَّرة لـ Y والموافقة للقيمة المعلومة x_o.

٤-٢-١-١- ملاحظة

إنَّ مستقيم الانحدار يمرُّ بالنقطة التي إحداثييها $(\bar{x}, \bar{y})$، ولذلك عندما نرغب برسم مستقيم الانحدار على لوحة الانتشار يكفي تعيين نقطة واحدة مختلفة عن النقطة $(\bar{x}, \bar{y})$ ومن ثمَّ الوصل بينهما بمستقيم.

٤-٢-٢-٣- أمثلة

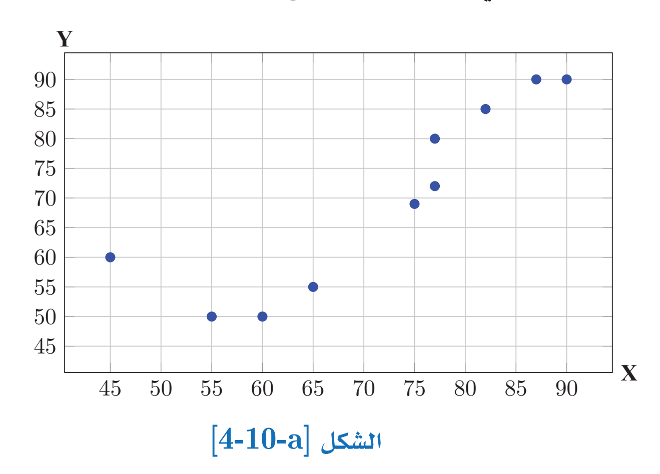
١- لتكن لدينا البيانات الآتية التي تمثّل درجات التحصيل النهائية لعشرة طلاب في مقرّري الرياضيات X والإحصاء Y:

الجدول [4–5-a]

رقم الطالب i	10	9	8	7	6	5	4	3	2	1
درجة الطالب في مقرَّر الرياضيات X	45	77	87	55	60	77	90	65	82	75
درجة الطالب في مقرَّر الإحصاء Y	60	72	90	50	50	80	90	55	85	69

ولنقم بتعيين معادلة مستقيم انحدار Y على X لهذه البيانات.

الحل: من أجل ذلك لنلقي نظرة على توضّع البيانات على لوحة الانتشار فنجد لها العرض الآتي.



الشكل [4-10-a]

فنلاحظ أنَّ مَنْحى هذه البيانات مُطَّرِداً، ولذلك سنتوقّع أن يكون لمستقيم الانحدار ميل موجب.

والآن لنستخدم الجدول الآتي من أجل تسهيل العمليات الحسابية في إنجاز صيغة العلاقة [4–2].

الجدول [4–5-b]

i	x_i	x_i^2	y_i	y_i^2	$x_i \cdot y_i$
1	75	5625	69	4761	5175
2	82	6724	85	7225	6970
3	65	4225	55	3025	3575
4	90	8100	90	8100	8100
5	77	5929	80	6400	6160
6	60	3600	50	2500	3000
7	55	3025	50	2500	2750
8	87	7569	90	8100	7830
9	77	5929	72	5184	5544
10	45	2025	60	3600	2700
المجموع	**713**	**52751**	**701**	**51395**	**51804**

ومنه يكون لدينا:

$$b = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} = \frac{10(51804) - (713 \times 701)}{10(52751) - (713)^2} = 0.952$$

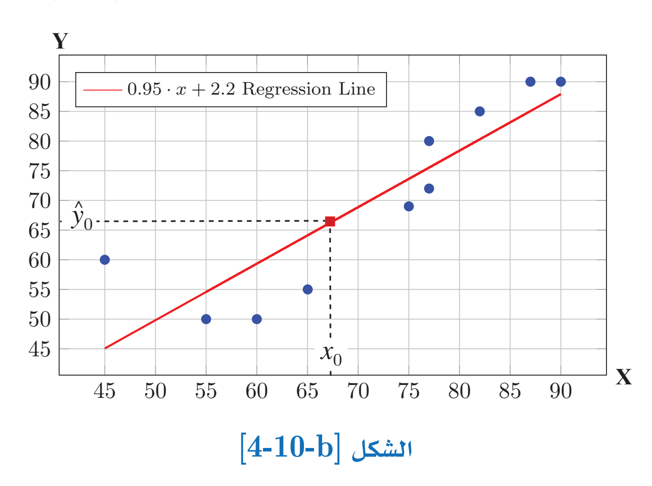
$$a=\frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i^2-\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n \sum_{i=1}^{n} x_i^2-\left(\sum_{i=1}^{n} x_i\right)^2}=\frac{701(52751)-(713\times 51804)}{10(52751)-(713)^2}=2.205$$

ومن ثمَّ يكون لمعادلة مستقيم انحدار Y على X العرض الآتي:

$$\hat{Y}=2.205+0.952\,X$$

حيث يُلاحظ بأنَّ ميله موجب (ويساوي 0.952) ويتقاطع مع المحور **oY** في النقطة $(0\ ,\ 2.205)$، فلو افترضنا أنَّ طالباً قد حصل على الدرجة $x_o=67.5$ في المقرَّر X، فعندئذ ستكون الدرجة المقدّرة له في المقرّر Y هي (انظر الشكل الآتي [4-10-b]):

$$\hat{y}_o=2.205+0.952(67.5)=66.465\approx 66.5$$



الشكل [4-10-b]

٢- البيانات الآتية تمثّل عيّنة من المبالغ التي دفعت من أجل المشتريات اليومية (مقدَّرةً بوحدة نقدية ما) من السلع الغذائية X والحاجات المنزلية الأخرى Y لأسرة.

الجدول [4-6-a]

i	1	2	3	4	5	6	7	8	9	10	11	12
X	13	33	100	85	16	45	75	31	42	77	36	85
Y	12	45	20	66	33	45	22	60	33	66	55	45

ولنقم بحساب مُعامل الارتباط الخطي، ومن ثمَّ تعيين معادلة مستقيم الانحدار الخطي لهذه البيانات.

الحل: من أجل ذلك سنقدِّم الحسابات الآتية.

الجدول [4–6-b]

i	x_i	x_i^2	y_i	y_i^2	$x_i \cdot y_i$
1	13	169	12	144	156
2	33	1089	45	2025	1485
3	100	10000	20	400	2000
4	85	7225	66	4356	5610
5	16	256	33	1089	528
6	45	2025	45	2025	2025
7	75	5625	22	484	1650
8	31	961	60	3600	1860
9	42	1764	33	1089	1386
10	77	5929	66	4356	5082
11	36	1296	55	3025	1980
12	85	7225	45	2025	3825
المجموع	**638**	**43564**	**502**	**24618**	**27587**

فنجد أنَّ:

$$\sum_{i=1}^{n} x_i = 638 \quad \& \quad \sum_{i=1}^{n} y_i = 502$$

$$\sum_{i=1}^{n} x_i^2 = 43564 \quad \& \quad \sum_{i=1}^{n} y_i^2 = 24618$$

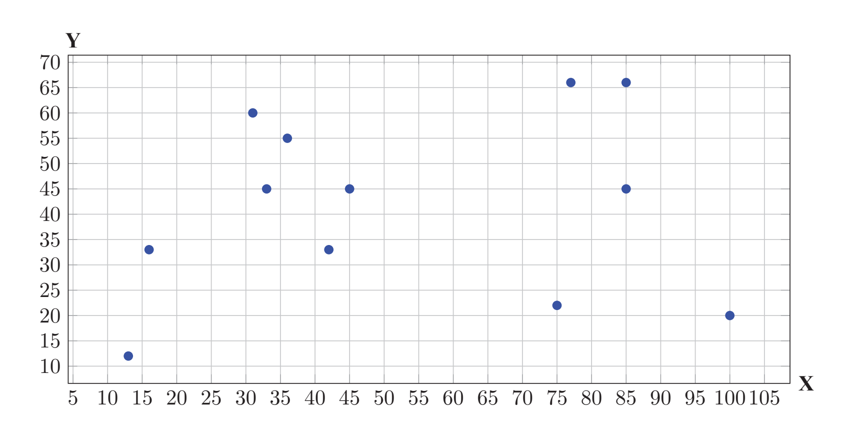
$$\sum_{i=1}^{n} x_i\, y_i = 27587 \quad \& \quad (\overline{x}, \overline{y}) = (53.17, 41.83)$$

ومنه تكون قيمة مُعامل الارتباط الخطي لهذه البيانات تساوي:

$$\mathbf{r} = \frac{n \cdot \left(\sum_{i=1}^{n} x_i y_i\right) - \left(\sum_{i=1}^{n} x_i\right) \cdot \left(\sum_{i=1}^{n} y_i\right)}{\sqrt{n \cdot \left(\sum_{i=1}^{n} x_i^2\right) - \left(\sum_{i=1}^{n} x_i\right)^2} \cdot \sqrt{n \cdot \left(\sum_{i=1}^{n} y_i^2\right) - \left(\sum_{i=1}^{n} y_i\right)^2}}$$

$$= \frac{12(27587) - (638 \times 502)}{\sqrt{12(43564) - (638)^2}\,\sqrt{12(24618) - (502)^2}} \approx 0.1519$$

وهذا يعني وجود ارتباط خطي **إيجابيّ ضعيف جداً** بين بيانات المتغيِّرين X وY، وهذا ما تأكّده لوحة الانتشار الآتية للبيانات المقدَّمة:



الشكل [4-11-a]

أمّا لتعيين مستقيم انحدار Y على X فإنَّنا نجد:

$$b = \frac{n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} = \frac{12(27587) - (638 \times 502)}{12(43564) - (638)^2} = 0.093$$

$$a = \frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} = \frac{502(43564) - (638 \times 27587)}{12(43564) - (638)^2} = 36.89$$

بالطبع كان من الممكن حساب الثابت a باستخدام العلاقة [4–4-b] لأنَّ قيم المتوسطات لـ X وY معلومة $(\bar{x}, \bar{y}) = (53.17, 41.83)$ فيكون لدينا:

$$a = \bar{y} - b\,\bar{x} = 41.83 - (0.093 \times 53.17) = 36.89$$

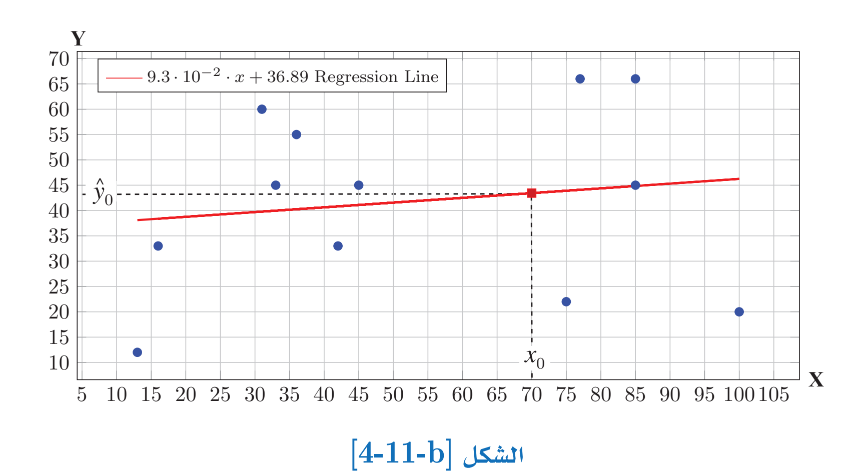
ومن ثمَّ يكون لمعادلة مستقيم انحدار Y على X العرض الآتي:

$$\hat{Y} = 36.89 + 0.093\,X$$

فلو أنّنا افترضنا أنَّ الأسرة قد اشترت سلع غذائية بمبلغ $x_o = 65$ وحدة نقدية، فعندئذ سيكون المبلغ المقدّر لشراء السلع الأخرى يساوي:

$$\hat{y}_o = 36.89 + (0.093 \times 65) = 42.935$$

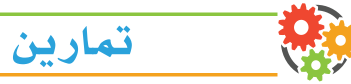
لاحظ التباعد الكبير للنقاط عن مستقيم الانحدار بسبب ضعف الارتباط بين بيانات المتغيِّرين X وY (انظر الشكل الآتي).


Y
$9.3 \cdot 10^{-2} \cdot x + 36.89$ Regression Line
70
65
60
55
50
45
40
35
30
25
20
15
10
$\hat{y}_0$
x_0
5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105
X

الشكل [4-11-b]



تمارين



١- اذكر أهم أنواع الارتباط من حيث التأثير والتأثر مع تقديم مثال على كل منها.

٢- ما الفائدة من استخدام لوحة الانتشار عند دراسة الارتباط؟

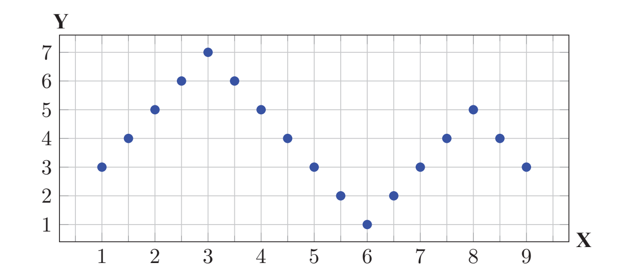
٣- لماذا ندرس الانحدار وما الفائدة منه؟

٤- لتكن لدينا مجموعة من البيانات الإحصائية الآتية:

i	1	2	3	4	5	6	7	8	9	10
X	1.2	1.3	1	1.5	1.1	1.3	2.5	2.5	0.5	0.7
Y	-1.6	-1.5	1.5	-1.2	0.3	1.1	0.5	0.7	0.3	- 0.3

والمطلوب تمثيل هذه البيانات على لوحة الانتشار.

٥- لتكن لدينا المشاهدات المقدّمة في الشكل الآتي:



والمطلوب ما يلي:

أ- بناء الجدول الخاص بهذه المشاهدات.

ب- حساب قيمة مُعامل الارتباط لهذه المشاهدات.

ج- تعيين معادلة مستقيم انحدار Y على X لهذه المشاهدات.

د- رسم مستقيم الانحدار على لوحة الانتشار المُعطاة.

٦- لتكن لدينا مجموعة من البيانات الإحصائية التي تُمثّل نتائج تسعة طلاب في اختبار المنتصف لمقرّري الإحصاء X والرياضيات Y.

i	1	2	3	4	5	6	7	8	9
X	30	20	19	28	07	25	29	18	27
Y	24	23	22	21	24	21	24	23	22

والمطلوب ما يلي:

أ- تمثيل هذه البيانات على لوحة الانتشار.

ب- حساب قيمة مُعامل الارتباط لهذه المشاهدات.

ج- تعيين معادلة مستقيم انحدار Y على X.

د- رسم مستقيم الانحدار على لوحة الانتشار.

هـ- استخدام مستقيم الانحدار لتقدير القيمة الموافقة لـ $x_o = 19.5$.

٧- أخضعت مجموعة مكونة من 10 رجال بدينين من ذوي الوزن 140 كغ لدورةٍ رياضيّةٍ ونظامٍ غذائيٍّ موحّدٍ من أجل تخفيض أوزانهم بحيث يمارس كل منهم فترة زمنية ما يومياً حسب استطاعته ولمدّة 45 يوماً، وبعد ذلك قيس الوزن لكل منهم، ومن ثمَّ دوِّن إلى جانب عدد الساعات التي مارسها كما في الجدول الآتي:

الرَجُل	X عدد الساعات التي مارسها خلال 45 يوماً	Y الوزن عند نهاية الدورة
1	30	131
2	45	125
3	56	122
4	42	128
5	90	103
6	33	131
7	48	120
8	35	127
9	70	110
10	82	108

عندئذ وباستخدام الارتباط والانحدار الخطي أجب على السؤال الآتي:

هل تشير هذه النتائج إلى درجةٍ عاليةٍ من التأثير المتبادل بين عدد ساعات الرياضة ونقصان الوزن؟

٨- لتكن لدينا المشاهدات الناتجة عن إخضاع مجموعة مكوَّنة من 12 طالباً لاختبارين تحريري X (الدرجة القصوى 80) ومقابلة Y (الدرجة القصوى 20) في أحد المقرَّرات الدراسيّة.

$(41,15)$, $(58,12)$, $(58,9)$, $(46,16)$, $(54,18)$, $(42,14)$,

$(64,13)$, $(60,16)$, $(29,5)$, $(71,19)$, $(79,18)$, $(62,17)$

والمطلوب ما يلي:

أ- تمثيل هذه البيانات على لوحة الانتشار.

ب- حساب قيمة مُعامل الارتباط الخطي لهذه المشاهدات.

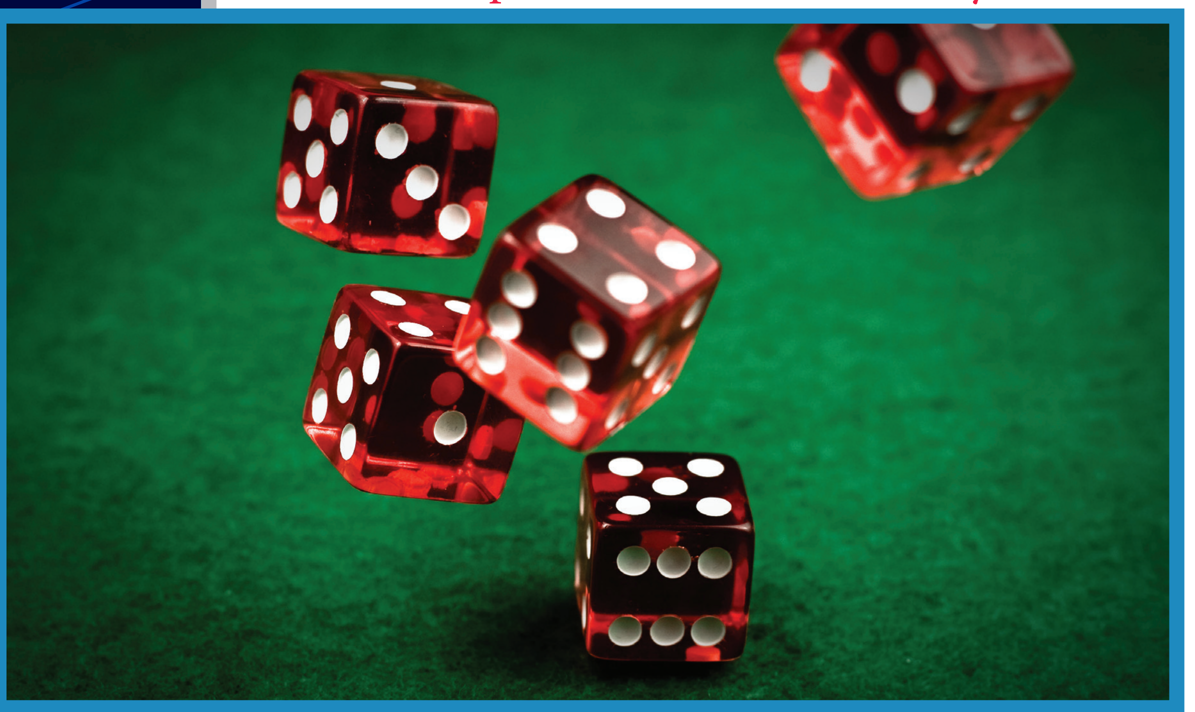
ج- تعيين معادلة مستقيم انحدار Y على X لهذه المشاهدات ومن ثمَّ رسمه على لوحة الانتشار.

هـ- استخدام معادلة مستقيم الانحدار لتقدير الدرجة $\hat{y}_o$ المقابلة $x_o = 68$.

الفصل الخامس

التجارب العشوائية واحتمالات الحوادث

Random Experiments and Probability of Events



المقدمة:

إنَّ مفهوم الحساب الاحتمالي لا يُعدّ من المفاهيم القديمة إذا ما قُورِنَ بمفهوم الإحصاء أو بمفاهيم رياضياتية أخرى مثل الجبر والهندسة و...، والحديث من هذا العلم لا يتجاوز عمره 85 سنةً حيث تبلور المفهوم الحديث للحساب الاحتمالي بعد وضع ما يُعرف باسم «مسلَّمات كلموغوراف» للفضاء الاحتمالي عام 1933 والتي قدّمها الرياضياتي الروسي كلموغوراف (1903-1987) Andrey Kolmogorov.

في الحقيقة لن نقوم بدراسة الاحتمالات وفقاً للنظرية الحديثة لهذا العلم وإنّما سنقدّمه وفق منهج تقليدي بسيط يتناسب وإمكانيات طلاب المسار الإنساني في الجامعات آخذين بالحسبان ما قد درسه الطالب من مواد علميّة في المراحل السابقة لتخصصه في الفرع الأدبي.

٥ - ١
القاعدتين الأساسيتين في العدّ

إنَّ من متطلبات الحساب الاحتمالي التعرّف على بعض قواعد العدّ والحساب التي يحتاجها الطالب في حساب بعض الاحتمالات لحوادث وسنبدأها بالمفهوم الآتي.

٥-١-١- قاعدة الضرب Multiplication Rule

لنفترض أنَّ عمليّة ما A يمكن أن تُنجز بـ n طريقة، وعمليّة أخرى B يمكن أن تُنجز بـ m طريقة، فعندئذٍ يمكننا إنجاز كلاً من العمليتين A و B بآنٍ واحدٍ بعددٍ من الطرائق يساوي $n \times m$. إنَّ هذا الإجراء الحسابي لهذا النوع من المسائل يُعرف باسم "قاعدة الضرب"، ويمكن تعميم هذه الطريقة في الحساب على أي عددٍ منتهٍ من العمليات التي لها عدد منتهٍ من الإمكانيات. فلو كان لدينا k عملية A_1 و A_2 و... و A_k بحيث يمكن أن تُنجز هذه العمليات بـ n_1 و n_2 و ... و n_k طريقة على الترتيب، فإنَّه يمكن إنجاز هذه العمليات جميعاً وبآن واحد بعدد من الطرائق يساوي $n_1 \times n_2 \times \ldots \times n_k$.

٥-١-١-١- أمثلة

١- لدينا قفل رقمي عشري لحقيبة مكوَّن من خمس خانات (كما في الشكل الجانبي).
بكم طريقة يمكننا أن نختار رقماً لقَفل هذا القفل؟

الحل: بملاحظة أن أول رقم في الخانة اليسرى يمكننا اختياره بعددٍ من الطرائق يساوي إلى 10، وكذلك الأمر بالنسبة إلى الخانات الأربع الأخرى المتبقية، فإنَّه يمكننا أن نختار رقماً لقفل هذا القفل بعدد من الطرائق يساوي:

$$10 \times 10 \times 10 \times 10 \times 10 = 100000$$

٢- إنَّ لوحات السيارات في بعض البلدان تُصمَّم بحيث تكون اللوحة المميِّزة للسيارة تحتوي على أربعة أرقام عشرية وثلاثة أحرف عربية، فما هو عدد السيارات التي يمكن أن تحمل لوحة من هذا النموذج؟

٢٠١٧	س أ م	KSA
2017	F Y C	

الحل: بملاحظة أنَّ أول رقم في اللوحة يمكننا اختياره بعددٍ من الطرائق يساوي إلى 10، وكذلك الأمر بالنسبة إلى بقية الأرقام الموجودة على اللوحة، أمّا بالنسبة إلى الأحرف فيمكننا اختيار أول حرف بعددٍ من الطرائق يساوي 28 طريقة، وبالمثل نجد من أجل اختيار بقية الأحرف الموجودة على اللوحة، ومن ثمَّ يكون عدد اللوحات التي يمكن إنتاجها وفقاً لهذا التقديم يساوي (وهو عدد السيارات التي يمكن أن تحمل لوحة من هذا النموذج):

$$10 \times 10 \times 10 \times 10 \times 28 \times 28 \times 28 = 219520000$$

٥-١-٢- قاعدة الجمع Addition Rule

لنفترض أنَّ عمليّة ما A يمكن أن تُنجز بـ n طريقة، وعمليّة أخرى B يمكن أن تُنجز بـ m طريقة، فعندئذٍ يمكن إنجاز العمليّة A أو العمليّة B بعدد من الطرائق يساوي $n+m$. إنَّ هذا الإجراء الحسابي لهذا النوع من المسائل يُعرف باسم "قاعدة الجمع"، ويمكن تعميم هذه التقنية في الحساب على أي عدد منته من العمليات التي لها عدد منته من الإمكانيات. فلو كان لدينا k عملية A_1 و A_2 و... و A_k بحيث يمكن أن تُنجز هذه العمليات بـ n_1 و n_2 و... و n_k طريقة على الترتيب، فإنَّه يمكن إنجاز العمليّة A_1 أو A_2 أو... أو A_k بعدد من الطرائق يساوي $n_1+n_2+\ldots+n_k$.

٥-١-١-١- أمثلة

١- يرغب أحد العاملين في أرشيف ترميز ملفاته باستخدام حرف عربي أو رقم عشري، فكم ملفاً يمكن ترميزه بهذه الطريقة؟

الحل: بملاحظة أنَّ العامل لا يستطيع استخدام إلّا حرف عربي أو رقم عشري فقط، فإنَّه يستطيع ترميز 28 ملفاً بوساطة الأحرف العربية و10 ملفات بوساطة الأرقام العشرية فقط، ومن ثمَّ يكون عدد الملفات التي يمكن ترميزها وفقاً لهذا التقديم يساوي $28+10=38$.

٢- يريد باحث تنفيذ تجربة تربوية على تلاميذ إحدى المدارس، حيث يمكنه تطبيقها إمَّا على التلاميذ الذكور ولديه ست مدارس فقط أو على التلميذات ولديه ثمانية مدارس. كم إمكانية لديه لتنفيذ تجربته؟

الحل: نلاحظ أنَّه يمكن للباحث تنفيذ تجربته بعدد من الطرائق يساوي $6+8=14$.

٣- تقوم مكتبة بترميز كتبها بوساطة إحدى الطريقتين الآتيتين:

- إمَّا باستخدام رقمين وثلاثة أحرف لاتينية،
- أو باستخدام ثلاثة أرقام وحرفين عربيين،

فما هو عدد الكتب التي يمكن ترميزها وفقاً لهذه الشروط؟

الحل: بملاحظة أنَّه يمكن للمكتبة أن ترمِّز باستخدام رقمين وثلاثة أحرف لاتينية عدداً من الكتب يساوي:

$$10\times10\times26\times26\times26=1757600$$

وكذلك باستخدام ثلاثة أرقام وحرفين عربيين تستطيع أن ترمِّز عدداً من الكتب يساوي:

$$10\times10\times10\times28\times28=784000$$

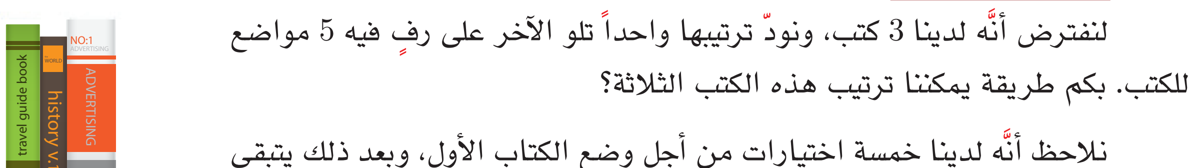
ومن ثمَّ يكون عدد الكتب التي يمكن ترميزها وفقاً لقاعدة الجمع يساوي:

$$1757600+784000=2541600$$

٥ - ٢
التراتيب والتوافيق

من المفاهيم الهامَّة في الحساب الاحتمالي التقليدي أيضاً ما يُعرف باسم "**التراتيب والتوافيق**"، والذين سنقدمهما من خلال الفقرتين الآتيتين:

٥-٢-١- التراتيب



لنفترض أنَّه لدينا 3 كتب، ونودّ ترتيبها واحداً تلو الآخر على رفٍ فيه 5 مواضع للكتب. بكم طريقة يمكننا ترتيب هذه الكتب الثلاثة؟

نلاحظ أنَّه لدينا خمسة اختيارات من أجل وضع الكتاب الأول، وبعد ذلك يتبقى لدينا أربعة مواضع شاغرة على الرف من أجل وضع الكتاب الثاني، وبعد وضع الكتاب الثاني يمكننا وضع الكتاب الأخير بإحدى ثلاث طرائق. إذاً، وبحسب قاعدة الضرب يمكننا أن نضع الكتب الثلاثة على هذا الرف واحداً تلو الآخر بعدد من الطرائق يساوي:

$$5 \cdot (5-1) \cdot (5-2) = 5 \times 4 \times 3 = \frac{5 \times 4 \times 3 \times 2 \times 1}{2 \times 1} = \frac{5!}{(5-3)!} = 60$$

في الحقيقة يمكننا تعميم هذه القاعدة من أجل عدد منتهٍ من الأشياء على النحو الآتي.

لنفترض أنَّه لدينا n من الأشياء المتمايزة ونريد أخذ k شيء منها واحداً تلو الآخر دون إعادة، فعندئذٍ يمكننا اختيار الشيء الأول بعدد من الطرائق يساوي n، وأمّا من أجل اختيار الشيء الثاني فيكون قد تبقى لدينا $n-1$ من إمكانيات الاختيار، ومن أجل اختيار الشيء الذي يليه يتبقى لدينا $n-2$ من إمكانيات الاختيار، وهكذا على هذا النحو فيتبقى لدينا من أجل اختيار الشيء الأخير $[n-(k-1)]$ من إمكانيات الاختيار، وبالتالي بحسب قاعدة الضرب يكون لدينا من أجل اختيار الـ k شيء عدداً من الطرائق يساوي:

$$n \cdot (n-1) \cdot \ldots \cdot [n-(k-1)] = \frac{n!}{(n-k)!}$$

ويُرمز لهذا المقدار بـ nPk، أي أنَّه لدينا:

$$nPk = n \cdot (n-1) \cdot \ldots \cdot [n-(k-1)] = \frac{n!}{(n-k)!} \qquad [5\text{-}1]$$

ويُقرأ "تراتيب لـ k شيء مأخوذة من n شيء متمايز على التتالي، علماً أنَّ:

$$n! = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 2 \cdot 1 \qquad [5\text{-}2]$$

ويُدعى " n مضروب **n-Factorial** "، والجدول الآتي يزودنا ببعض القيم لـ $n!$ من أجل استخدامها في حل التمارين والمسائل.

الجدول [5-1]

n	0	1	2	3	4	5	6	7	8	9	10
$n!$	1	1	2	6	24	120	720	5040	40320	362880	3628800

وأمّا إذا كانت قيمة n كبيرة نسبياً فمن المفيد استخدام العلاقة الآتية التي تنتج من نشر سترلنغ Stirling Expansion (نسبة إلى الرياضياتي الأسكتلندي سترلنغ **James Stirling (1692-1770)**) لحساب المقدار $n!$:

$$n! \approx \sqrt{2\, n\, \pi}\;\; n^n \;\; \mathbf{e}^{-n} \qquad ; n \in \mathbb{N}$$

الآن، وفي الحالة الخاصَّة عندما يصبح لدينا $k = n$ فإنَّه يُقال عن nPn تبديل لـ n من الأشياء المتمايزة، ويكون لدينا في هذه الحالة:

$$nPn = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 2 \cdot 1 = n\,!$$

سنقدِّم فيما يلي أمثلة بسيطة على سبيل التوضيح.

◀ ٥-١-٢-١- مثال

١- لدينا صندوق يحوي 3 كرات متماثلة تماماً ومرقّمة بالأعداد 1، 2 و3. عندئذٍ:

أ- إذا قمنا بسحب عشوائي لكرتين من الصندوق على التوالي وبدون إعادة، فبكم طريقة يمكننا اختيار هاتين الكرتين؟

ب- إذا قمنا بسحب عشوائي لثلاث كرات من الصندوق على التوالي وبدون إعادة، فبكم طريقة يمكننا اختيار هذه الكرات الثلاث؟

ج- نسحب عشوائياً كرتين من الصندوق على التوالي ومع الإعادة، فبكم طريقة يمكننا اختيار هاتين الكرتين؟

الحل: من أجل الطلب:

أ- بملاحظة أنَّه يمكننا سحب الكرة الأولى بعدد من الطرائق يساوي 3، والتي تليها بعدد من الطرائق يساوي 2، ومن ثمَّ بحسب قاعدة الضرب يمكننا اختيار الـكرتين بعدد من الطرائق يساوي:

$$3 \times 2 = 6 = \frac{3!}{(3-2)!}$$

وهذا موافق للسحوبات الممكنة الآتية حيث يُلاحظ أنَّ للترتيب أهمية في هذا النوع من السحب:

$(1,2),\ (1,3),\ (2,1),\ (2,3),\ (3,1),\ (3,2)$

ب- بملاحظة أنَّه يمكننا سحب الكرة الأولى بعدد من الطرائق يساوي 3، والتي تليها بعدد من الطرائق يساوي 2، والكرة الأخيرة بعدد من الطرائق يساوي 1، ومن ثمَّ بحسب قاعدة الضرب يمكننا اختيار الكرات الثلاث بعدد من الطرائق يساوي:

$$3 \times 2 \times 1 = \frac{3!}{(3-3)!} = \frac{3!}{0!} = 6 = 3P3$$

وهذا يعني أنَّه لدينا في هذه الحالة تباديل لثلاثة أشياء. لاحظ أنَّ للترتيب أهمية في هذا النوع من السحب أيضاً حيث لدينا:

$$(1,2,3),\ (1,3,2),\ (2,1,3),\ (2,3,1),\ (3,1,2),\ (3,2,1)$$

ج- بملاحظة أنَّه يمكننا سحب الكرة الأولى بعدد من الطرائق يساوي 3، وكذلك التي تليها يمكننا سحبها بعدد من الطرائق يساوي 3 أيضاً، ومن ثمَّ بحسب قاعدة الضرب يمكننا اختيار الكرتين بعدد من الطرائق يساوي:

$$3 \times 3 = 3^2 = 9$$

وهذا موافق للسحوبات الممكنة الآتية حيث يُلاحظ أنَّ للترتيب أهمية في هذا النوع من السحب:

$$(1,1),\ (1,2),\ (1,3),\ (2,1),\ (2,2),\ (2,3),\ (3,1),\ (3,2),\ (3,3)$$

ولكن لم تعد المسألة هي مسألة تراتيب رغم أهمية الترتيب لنتائج السحب، والسبب في ذلك يعود إلى عملية الإعادة للكرات المسحوبة.

▶

٥-٢-٢- التوافيق

لنفترض أنَّه لدينا n من الأشياء المتمايزة ونريد أخذ k شيء منها دفعة واحدة، فعندئذٍ يكون لدينا عدداً من الطرائق لاختيار الـ k شيء يساوي $\frac{n!}{k!\,(n-k)!}$، ويُدعى هذا المقدار توافيق لـ n فوق k، ويُرمز له بـ nCk أو بـ $\binom{n}{k}$. أي أنَّه لدينا:

$$nCk = \binom{n}{k} := \frac{n!}{k!\,(n-k)!} \qquad [5\text{-}3]$$

٥-٢-٢-١- ملاحظة

في الواقع يمكن طرح المسألة السابقة على وجه آخر، فهي تكافئ مسألة أخذ k شيءٍ واحدٍ تلو الأخرى من n شيءٍ متمايزٍ وبدون إعادةٍ، ومع الأخذ بالحسبان أنَّه لا أهمية لترتيب الأشياء المسحوبة، فعلى سبيل المثال لو كانت هذه الأشياء مرقّمة من 1 وحتى n، فعندئذٍ إذا حصلنا في السحب الأول على الشيء الذي يحمل الرقم 5 وفي السحب الثاني على الشيء الذي يحمل الرقم 9 هو نفسه كما لو حصلنا على الشيء الذي يحمل الرقم 9 في السحب الأول وعلى الشيء الذي يحمل الرقم 5 في السحب الثاني، والمهمّ في هذه الحالة أنَّنا حصلنا على 5 و9 في السحبين الأول والثاني.

على هذا النحو تُناقش بقية السحوبات الممكنة، ولذلك سيكون لدينا عدد الطرائق الممكنة لسحب k شيءٍ من n شيءٍ متمايز يساوي:

$$\frac{n\cdot(n-1)\cdot(n-2)\cdot\ldots\cdot n-(k-1)}{k!}=\frac{n!}{k!\,(n-k)!}=nCk$$

والأمثلة الآتية توضِّح لنا ذلك.

◄ ٥-٢-٢-٢- أمثلة

١- لدينا صندوق يحوي 3 كرة متماثلة تماماً ومرقّمة بالأعداد 1 وحتى 3. نسحب عشوائياً كرتين من الصندوق دفعة واحدة. بكم طريقة يمكننا اختيار هاتين الكرتين؟

الحل: بملاحظة أنَّ السحوبات الممكنة هي $(1,2)$، $(1,3)$ و $(2,3)$، وذلك لأنَّ النتيجتين $(1,2)$ و $(2,1)$ تعدّ نتيجة واحدة فقط بسبب عدم أهمية الترتيب، وذلك لأنَّنا لا نعلم أي منهما الأولى وأيها الثانية بسبب سحبهما بآن واحد. كذلك الأمر بالنسبة إلى النتيجتين $(1,3)$ و $(3,1)$ والنتيجتين $(2,3)$ و $(3,2)$.

إذاً، فعدد الطرائق الممكنة لسحب هاتين الكرتين يساوي $3C2$ حيث لدينا:

$$3C2=\frac{3!}{2!\,(3-2)!}=\frac{3!}{2!\cdot 1!}=\frac{6}{(2)\cdot(1)}=3$$

٢- في مركز صحي يوجد تسع ممرضات وأربعة أطباء، وطُلِب منّا تشكيل لجنةٍ من أجل إدارة هذا المركز بحيث تكون مكوَّنة من طبيبين وثلاث ممرضات. بكم طريقة يمكننا تشكيل هذه اللجنة؟

الحل: بملاحظة أنَّه يمكننا أن نختار طبيبين من أربعة أطباء بعدد من الطرائق يساوي $4C2$ (لأنَّه لا أهمية للترتيب هنا، فالمهم الحصول على طبيبين فقط)، وكذلك يمكننا أن نختار ثلاث ممرضات من تسع ممرضات بعدد من الطرائق يساوي $9C3$ (لأنَّه لا أهمية للترتيب هنا أيضاً، والمهم الحصول على ثلاث ممرضات)، ومن ثمَّ بحسب قاعدة الضرب يمكننا اختيار طبيبين وثلاث ممرضات بعدد من الطرائق يساوي $(4C2)\cdot(9C3)$، حيث لدينا:

$$(4C2)\cdot(9C3)=\frac{4!}{2!\,(4-2)!}\cdot\frac{9!}{3!\,(9-3)!}=6\times 84=504$$

٣- شَهِد وقوع حادثة جنائية في مركز للتسوق سبعة رجال وثمان نساء. بكم طريقة يمكن اختيار الشهود على هذه الواقعة (وفقاً للقاعدة الشرعية للشهادة)؟

الحل: من المعلوم أنَّ الشهادة تتمّ برجلين أو رجل وامرأتين، وبما أنَّ الترتيب بين كل الرجال والنساء من الشهود ليس له أهمية فإنَّه يمكننا اختيار الشهود على هذه الواقعة بعدد من الطرائق يساوي (مستخدمين في ذلك قاعدتي الجمع والضرب):

$$7C2+\left(7C1\right)\cdot\left(8C2\right)=\frac{7!}{2!(7-2)!}+\frac{7!}{1!(7-1)!}\cdot\frac{8!}{2!(8-2)!}$$
$$=21+(7\times 28)=217$$

٥ - ٣
فضاء الحوادث الابتدائية

إنَّ مفهوم فضاء الحوادث الابتدائية ذو صلة وطيدة بمفهوم التجارب العشوائية، ولذلك لا بدَّ لنا أولاً من توضيح مفهوم التجربة العشوائية.

في الواقع إنَّ التجارب التي يقوم بها المرء تقسم إلى نوعين:

النوع الأول: تجارب نتائجها معروفة مسبقاً وقابلة للتخمين بشكل دقيق، والأمثلة على ذلك كثيرةٌ جداً، ومنها على سبيل المثال لا الحصر تجربة وضع ملعقة من السكر في كوب شاي ومن ثمَّ تحريك السائل حيث تنحل بلورات السكر في الماء ونحصل على كوب من الشاي المُحلَّى، ويمكن للكيميائيين تحديد درجة حلاوة السائل في الكوب (وبدقّة) مسبقاً إذا علموا كمية المقادير التي استخدمت مع خصائصها. إنَّ هذا النوع من التجارب يُدعى **تجارب نظامية** Regular Experiments.

النوع الثّاني: تجارب لا يمكن الجزم بمعرفة نتائجها مسبقاً، وكل ما يُمكن عمله تعيين مجموعة تنتمي إليها تلك النتائج، والأمثلة على ذلك كثيرة أيضاً، ومنها على سبيل المثال لا الحصر تجربة قذف قطعة نقود معدنية Coin لمرَّةٍ واحدة، فعندئذٍ سنحصل إمّا على صورة H أو على شعار T، ولكن أي من هاتين النتيجتين ستظهر للأعلى؟

بالطبع لا يمكن الجزم بمعرفتها مسبقاً وكل ما يمكننا فعله في الواقع هو تعيين مجموعة تنتمي إليها نتائج هذه التجربة ألا وهي $\{H, T\}$.

إنَّ هذا الصنف من التجارب يُدعى **تجارب عشوائية** Random Experiments.

٥-٣-١- الحادث الابتدائي وفضاء الحوادث الابتدائية

في الواقع ينتج عن كل تجربة عشوائية مجموعة معيّنة (أو محدَّدة) من النتائج الممكنة يُرمز لها عادةً بـ Ω . إنَّ كل نتيجة من هذه النتائج تُدعى **حادثاً ابتدائياً** Elementary Event، ولهذا السبب يُطلق على مجموعة كل نتائج التجربة العشوائية Ω اسم **"فضاء الحوادث الابتدائية"** Space of Elementary Events. إنَّ تعيين فضاء الحوادث الابتدائية Ω لتجربة عشوائية يحتاج إلى انتباهٍ ودقةٍ كبيرين في فهم التجربة العشوائية وذلك لأنَّه لا توجد قاعدة مُحدَّدة من أجل استنباط مجموعة النتائج لتجربة عشوائية.

فيما يلي نقدِّم مجموعة من الأمثلة البسيطة نوضِّح من خلالها كيفية تعيين مجموعة نتائج تجربة عشوائية (أو فضاء الحوادث الابتدائية)، ولكن من أجل تجارب عشوائية ذات عددٍ منتهٍ من النتائج فقط.

٤-٣-٢- أمثلة

١- لنقذف قطعة نقود معدنية لمرّتين متتاليتين، فعندئذٍ ستكون النتائج كما في العروض الآتية:

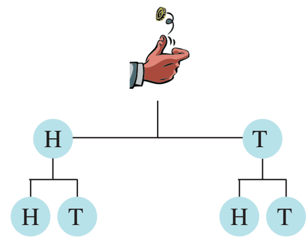


وهنا نلاحظ أنَّ نتائج هذه التجربة هي ثنائيات مُركّباتها إمّا شعاراً أو صورةً، ومن ثمَّ يكون لمجموعة نتائج هذه التجربة العشوائية العرض الآتي:

$$\Omega = \{(H,H), (H,T), (T,H), (T,T)\}$$

حيث نجد أنَّ عدد النتائج الممكنة لهذه التجربة العشوائية منتهٍ، ويساوي $|\Omega| = 4 = 2^2$.

ننوّه هنا إلى أنَّه يمكن تمثيل نتائج هذه التجربة بشكل بياني يُدعى **تمثيل الشجرة** ويوضّحه الشكل الآتي [5-1].



الشكل [5-1]

كذلك نشير إلى أنَّنا سنكتب (على سبيل التبسيط ومالم يؤدي ذلك إلى التباس) المجموعة السابقة Ω على النحو الآتي:

$$\Omega = \{HH, HT, TH, TT\}$$

وبالمثل بالنسبة إلى المجموعات التي تماثلها من حيث العرض.

٢- لنقذف قطعتي نقود **متمايزتين** (مختلفتين) لمرّة واحدة فقط، فعندئذٍ ستكون نتائج التجربة هي عبارة عن ثنائيات أيضاً، ومُركّباتها إمّا شعاراً أو صورةً كما في العروض الآتية (في احدى قطعتي النقود نقشت كتابة بدلاً من الصورة):



ومن ثمَّ تكون مجموعة النتائج لهذه التجربة العشوائية هي:

$$\Omega = \{HH, HT, TH, TT\}$$

وهنا نلاحظ أنَّ النتائج الممكنة لهذه التجربة العشوائية متطابقة مع النتائج الممكنة للتجربة في المثال السابق، والسبب في ذلك يعود لتمايز قطعتي النقود.

٣- لنقم بقذف قطعتي نقود غير متمايزتين (متماثلتين تماماً) بآنٍ واحد ولمرّة واحدة فقط، فعندئذٍ ستكون نتائج هذه التجربة هي ثنائيات مُركّباتها شعار أو صورة كما في العروض الآتية:



وذلك لأنّ عدم التمايز لقطعتي النقود يجعلنا ننظر إلى النتيجتين:



على أنّها نتيجة واحد فقط، وذلك لأنّه لا يمكننا معرفة تبعية المُركّبة الأولى (وبالمثل المُركّبة الثانية) أهي من القطعة الأولى أم الثانية، ومن ثمّ يكون لمجموعة نتائج هذه التجربة العرض الآتي:

$$\Omega = \{HH, HT, TT\}$$

حيث نجد أنّ عدد النتائج الممكنة لهذه التجربة العشوائية يساوي $|\Omega| = 3$

٤- لنأخذ تجربة رمي حجر نرد لمرّة واحدة فقط، فعندئذٍ ستكون مجموعة النتائج لهذه التجربة العشوائية هي:

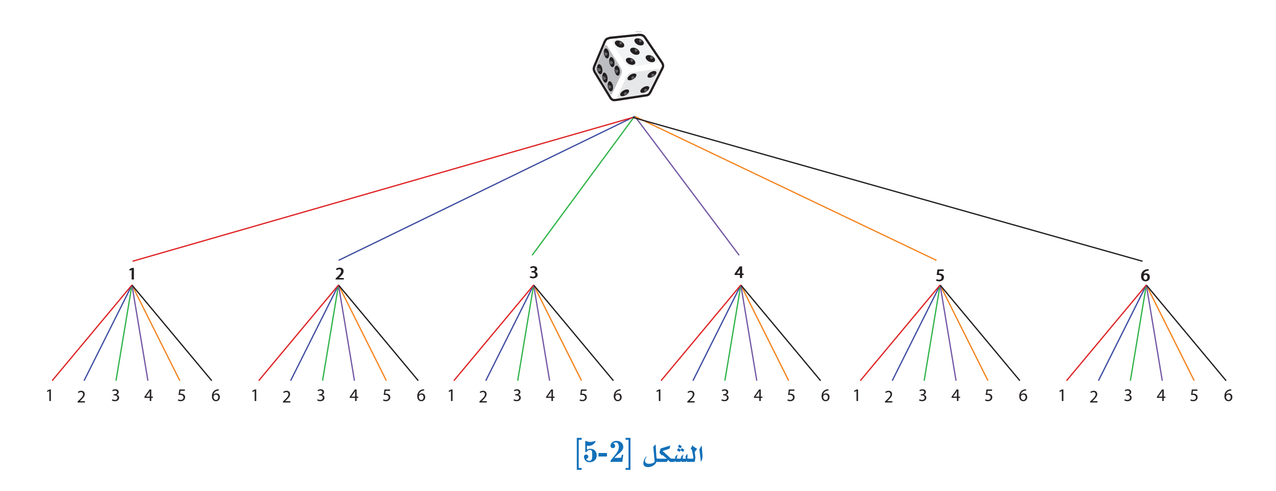
$$\Omega = \{1,2,3,4,5,6\}$$

فنلاحظ أنّ عدد النتائج الممكنة لهذه التجربة العشوائية هو $|\Omega| = 6$.

٥- لنأخذ تجربة رمي حجر نرد لمرّتين متتاليتين، فعندئذٍ سيكون لمجموعة نتائج هذه التجربة العشوائية العرض الآتي (وتمثيل الشجرة لنتائج هذه التجربة العشوائية موضّح أدناه):

$$\Omega = \begin{Bmatrix} (1,1),(1,2),(1,3),(1,4),(1,5),(1,6), \\ (2,1),(2,2),(2,3),(2,4),(2,5),(2,6), \\ (3,1),(3,2),(3,3),(3,4),(3,5),(3,6), \\ (4,1),(4,2),(4,3),(4,4),(4,5),(4,6), \\ (5,1),(5,2),(5,3),(5,4),(5,5),(5,6), \\ (6,1),(6,2),(6,3),(6,4),(6,5),(6,6) \end{Bmatrix}$$

فنجد أنّ عدد النتائج الممكنة لهذه التجربة العشوائية هو $|\Omega| = 36 = 6^2$، وأمّا عرض الشجرة لنتائج هذه التجربة العشوائية فيقدّمه الشكل الآتي:



الشكل [5-2]

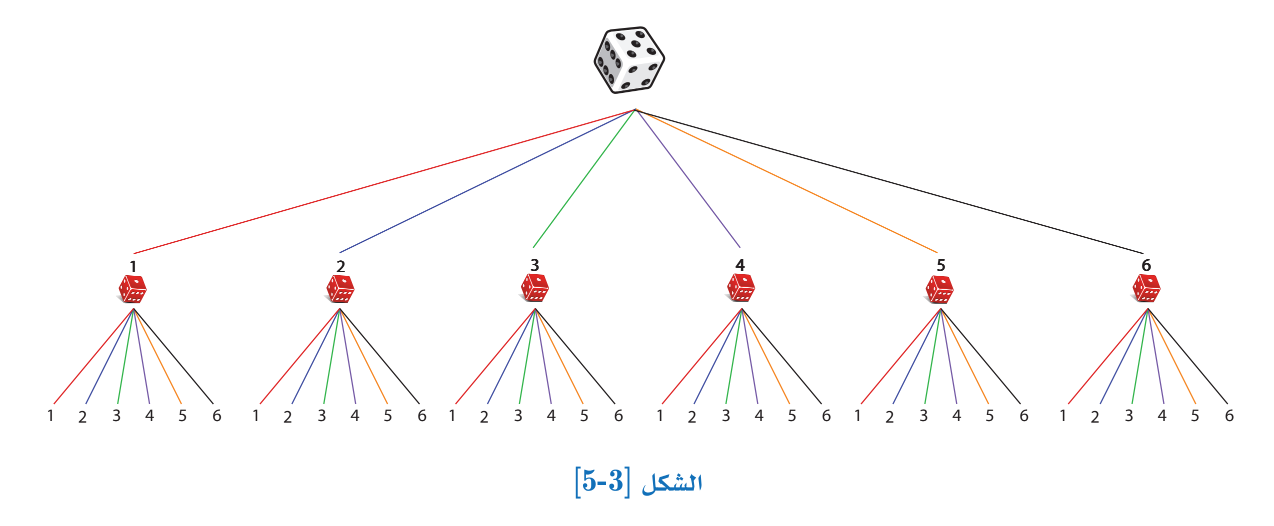


٦- لنأخذ تجربة رمي حجري نرد **متمايزين** بآنٍ واحد ولمرّة واحدة فقط، فعندئذٍ ستكون لمجموعة نتائج هذه التجربة العشوائية العرض الآتي (وتمثيل الشجرة لنتائج هذه التجربة العشوائية موضّح في الشكل [5-3]):

$$\Omega=\left\{\begin{matrix}(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),\\(2,1),(2,2),(2,3),(2,4),(2,5),(2,6),\\(3,1),(3,2),(3,3),(3,4),(3,5),(3,6),\\(4,1),(4,2),(4,3),(4,4),(4,5),(4,6),\\(5,1),(5,2),(5,3),(5,4),(5,5),(5,6),\\(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\end{matrix}\right\}$$

فنجد أنَّ عدد النتائج الممكنة لهذه التجربة العشوائية هو $6^2 = 36 = |\Omega|$ أيضاً، والسبب في ذلك يعود لتمايز حجري النرد.

أمَّا عرض الشجرة لنتائج هذه التجربة العشوائية فهو مماثل للشكل [5-2] ويقدِّمه الشكل الآتي:



الشكل [5-3]

٥ - ٤
الحوادث

من المعلوم أنَّه من أجل تجربة عشوائية ما يتوجَّب على المرء البت فيما إذا كانت النتيجة المطلوبة قد تحقَّقت أم لا، وذلك من أجل مقولة مُحدّدة متعلّقة بهذه التجربة، فعلى سبيل المثال لدى تجربة إلقاء حجر نرد لمرَّةٍ واحدةٍ فقط قد نكون مهتمين بحصولنا على عدد فردي، ومن ثمَّ علينا البت فيما إذا كانت نتيجة هذه التجربة قد تحقَّقت من أجل المقولة التي ذكرناها (الحصول على عدد فردي) أم لا. في الحقيقة إنَّ هذا الطرح يكافئ القول الآتي:

إنَّ المجرِّب لا يهتمُّ عادةً بأية نتيجة للتجربة سيأخذ، وإنَّما الذي يهمُّه هو إن كانت هذه النتيجة ستنتمي إلى مجموعة جزئية أو أكثر من مجموعة كل نتائج التجربة العشوائية أم لا.

فيما يلي سنقدِّم مفهوم الحادث من أجل تجربة عشوائية منتهية النتائج فقط، وذلك لأنَّه من أجل الحالات التي تكون فيها مجموعة نتائج التجربة العشوائية غير منتهية (قابلة للعدّ أو غير قابلة للعدّ) تحتاج لمفاهيم رياضياتية تقع خارج إطار هذا الكتاب.

٥-٤-١- تعريف الحادث Event

لتكن لدينا تجربة عشوائية مجموعة نتائجها Ω منتهية. عندئذٍ كل مجموعة جزئية A من Ω تُدعى **حادثاً** Event.

٥-٤-٢- حوادث مستنتجة

لتكن لدينا تجربة عشوائية مجموعة نتائجها $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ مع n عدد طبيعي كيفي ولكن مثبّت، ولنفترض أنَّ A وB حادثين متعلّقين بهذه التجربة، فعندئذ:

١- إذا كان الحادث A يحوي حادث ابتدائي وحيد، أي أنَّ $A = \{\omega_i\}$ مع i قيمة ما من $\{1, 2, \dots, n\}$، فعندئذٍ يُقال عن A إنَّه **حادث بسيط**.

٢- حادث تَحقُّق A **أو** B (ويُرمز له بـ $A \cup B$) هو ذلك الحادث الذي يُعبِّر عن انتماء نتيجة التجربة إلى A أو B أو إلى كليهما.

٣- حادث تَحقُّق A **و** B معاً (ويُرمز له بـ $A \cap B$) هو ذلك الحادث الذي يُعبِّر عن انتماء نتيجة التجربة إلى كلٍ من A و B بآنٍ واحد.

٤- حادث تَحقُّق A ولكن دون B (ويُرمز له بـ $A \setminus B$)، هو ذلك الحادث الذي يُعبِّر عن انتماء نتيجة التجربة إلى الحادث A ولكن دون أن تنتمي إلى الحادث B.

٥- الحادث الذي يتحقَّق في حال عدم تحقُّق الحادث A يُدعى حادثاً متمِّماً Complement Event للحادث A، وسنرمز له بـ $\overline{A}$، أي أنَّ $\overline{A} = \Omega \setminus A$.

٦- من أجل أي حادث $A \subseteq \Omega$ سيكون من المستحيل انتماء نتيجة التجربة إلى كلٍ من الحادث A ومتممّه $\overline{A}$ بآنٍ واحد، ولذلك يُسمّى الحادث $A \cap \overline{A}$ بـ الحادث المستحيل Impossible Event، وبما أنَّ $A \cap \overline{A} = \varnothing$ (وفقاً لنظرية المجموعات) فإنَّه يُرمز للحادث المستحيل بـ$\varnothing$ أيضاً.

٧- من أجل أي حادث $A \subseteq \Omega$ ستنتمي نتيجة التجربة إلى الحادث $A \cup \overline{A}$ بكل تأكيد، ولذلك يُسمّى الحادث $A \cup \overline{A}$ بـ الحادث الأكيد Certain Event، وبما أنَّ $\Omega = A \cup \overline{A}$ (وفقاً لنظرية المجموعات) فإنَّه يُرمز للحادث الأكيد بـ Ω أيضاً.

٨- إذا كان انتماء نتيجة التجربة إلى الحادثين A وB بآنٍ واحد أمراً غير ممكن (مستحيلاً)، فحينئذ يُقال إنَّ A وB حادثين متنافيين Exclusive Events، حيث يكون لدينا في هذه الحالة $A \cap B = \varnothing$.

٩- إذا كانت A_1 و A_2 و... و A_k حوادث من Ω، وكان $A_i \underset{i \neq j}{\cap} A_j = \varnothing$ من أجل كل القيم الممكنة لـ i وj من $\{1,2,...,k\}$، فحينئذ يُقال عن هذه الحوادث إنَّها متنافيةً مثنى مثنى Pairwise Mutually Exclusive.

٥-٤-٣- ملاحظات

١- نشير هنا إلى أنَّه عندما نذكر مستقبلاً أنَّه في تجربة عشوائية ما (مثل: قذف قطع نقود معدنية أو رمي حجارة نرد أو سحب كرات من صندوق أو...) أنَّ الأدوات المستخدمة في توليد النتائج (مثل: قطع النقود المعدنية أو حجارة النرد أو الكرات التي في صندوق أو...) هي أدوات متماثلة إنَّما نَقصد بذلك أنَّه لا يمكننا التمييز بين هذه الأدوات، وفي حال ذكرنا أنَّ هذه الأدوات متمايزة إنَّما نَقصد بذلك أنَّ النتائج المولَّدة بتلك الأدوات يمكن تمييز بعضها عن البعض الآخر.

٢- سنرمز بـ 2^{Ω} لأسرة كل المجموعات الجزئية في Ω، فإذا كانت Ω منتهية مع $|\Omega| = n$ فعندئذٍ سيكون عدد العناصر في 2^{Ω} يساوي 2^n، ولتوضيح ذلك سنقدّم الأمثلة الآتية:

أ- بفرض أنَّ $\Omega = \{H, T\}$، فعندئذٍ يكون لأسرة كل المجموعات الجزئية في المجموعة Ω العرض الآتي:

$$2^{\Omega} = \left\{ \varnothing, \{H\}, \{T\}, \{H, T\} = \Omega \right\}$$

ومن ثمَّ تكون $\left|2^{\Omega}\right| = 2^2 = 4$.

ب- بفرض أنَّ $\Omega = \{3,4,5\}$ هي فضاء الحوادث الابتدائية لتدوير مثلث قائم (أضلاعه 3 ، 4 و 5 وحدات قياس) حول مركز ثقله ورصد طول الضلع الذي سيظهر للأعلى، فعندئذٍ يكون لأسرة كل المجموعات الجزئية في المجموعة Ω العرض الآتي:

$$2^{\Omega} = \left\{ \varnothing, \{3\}, \{4\}, \{5\}, \{3,4\}, \{3,5\}, \{4,5\}, \{3,4,5\} = \Omega \right\}$$

فنجد أنَّ:

$$\left|2^{\Omega}\right| = 2^3 = 8$$

ج- بفرض أنَّ $\Omega = \{HH, HT, TH, TT\}$، فعندئذٍ يكون لأسرة كل المجموعات الجزئية في المجموعة Ω العرض الآتي:

$$2^{\Omega} = \left\{ \begin{array}{l} \varnothing, \{HH\}, \{HT\}, \{TH\}, \{TT\}, \{HH,HT\}, \{HH,TH\}, \\ \{HH,TT\}, \{HT,TH\}, \{HT,TT\}, \{TH,TT\}, \{HH,HT,TH\}, \\ \{HH,HT,TT\}, \{HH,TH,TT\}, \{HT,TH,TT\}, \\ \{HH,HT,TH,TT\} = \Omega \end{array} \right\}$$

ومنه يكون لدينا:

$$\left|2^{\Omega}\right| = 2^4 = 16$$

◄ ٥-٤-٤- أمثلة

١- لنأخذ تجربة رمي حجر نرد لمرّةٍ واحدةٍ فقط، ولنأخذ:

الحادث A، والذي يُعبِّر عن حصولنا على عدد أصغر من 3،

الحادث B، والذي يُعبِّر عن حصولنا على عدد أكبر من 4،

الحادث C، والذي يُعبِّر عن حصولنا على عدد أكبر أو يساوي 3 وأصغر أو يساوي 4،

الحادث D، والذي يُعبِّر عن حصولنا على عدد أصغر من 1،

الحادث E، والذي يُعبِّر عن حصولنا على عدد أكبر أو يساوي 1.

الحادث F، والذي يُعبِّر عن حصولنا على عدد فردي.

الحادث G، والذي يُعبِّر عن حصولنا على عدد زوجي.

عندئذٍ نجد أنَّ الحوادث A وB وC هي:

$$A = \{1,2\} \quad \& \quad B = \{5,6\} \quad \& \quad C = \{3,4\}$$

وهذه الحوادث متنافية مثنى مثنى وذلك لأنَّ:

$$A \cap B = \{1,2\} \cap \{5,6\} = \varnothing$$

$$A \cap C = \{1,2\} \cap \{3,4\} = \varnothing$$
$$B \cap C = \{5,6\} \cap \{3,4\} = \varnothing$$

في حين نجد أنَّ الحادث D هو الحادث المستحيل لأنَّ $D = \{\ \}$، وأمّا الحادث E فهو الحادث الأكيد حيث لدينا $E = \{1,2,3,4,5,6\} = \Omega$، وأخيراً نجد أنَّ الحادثين F وG هما:

$$F = \{1,3,5\} \qquad \& \qquad G = \{2,4,6\}$$

وهما متنافيان وذلك لأنَّ:

$$F \cap G = \{1,3,5\} \cap \{2,4,6\} = \varnothing$$

لاحظ هنا أنَّ:

- الحادثين F وA وكذلك F وB وأخيراً F وC ليست متنافية.
- الحادثين G وA وكذلك G وB وأخيراً G وC ليست متنافية أيضاً.

٢- لنفترض أنَّه لدينا صندوق يحوي 4 كرات متماثلة تماماً، ولكل كرتين منها لون مميّز (أحمر R وأزرق B)، وقد رقِّمت كل كرة من الكرتين ذات اللون الواحد بالرقم 1 و2 للتمييز. نقوم بسحب عشوائي (أي بعد خلط الكرات جيداً ودون النظر إلى الصندوق أثناء السحب) لكرتين على التتالي دون إرجاع إلى الصندوق، والمطلوب تعيين:



أ- حادث حصولنا على كرتين من لون واحد.

ب- حادث حصولنا على أرقام مختلفة.

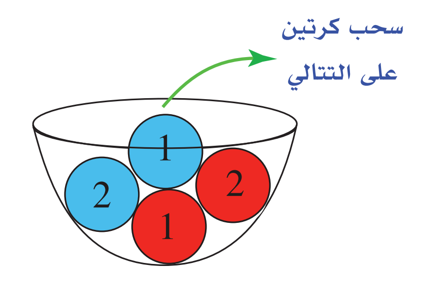
ج- حادث حصولنا على مجموع للأرقام يساوي 4.

د- حادث حصولنا على مجموع للأرقام يساوي 3.

هـ- حادث حصولنا على مجموع للأرقام أصغر من 5.

و- حادث حصولنا على مجموع للأرقام أصغر من 2.

ز- حادث حصولنا على كرات تحمل أرقاماً فردية.



الحلول: من أجل مسألتنا هذه سنرمز بـ r_1 و r_2 للكرة الحمراء التي تحمل الرقم 1 و2 على الترتيب، وكذلك بـ b_1 و b_2 للكرة الزرقاء التي تحمل الرقم 1 و2 على الترتيب، فعندئذٍ يكون لفضاء الحوادث الابتدائية العرض الآتي:

$$\Omega = \left\{ \begin{array}{l} (r_1,r_2),(r_2,r_1),(b_1,b_2),(b_2,b_1),(r_1,b_1),(r_1,b_2), \\ (r_2,b_1),(r_2,b_2),(b_1,r_1),(b_1,r_2),(b_2,r_1),(b_2,r_2) \end{array} \right\}$$

والآن من أجل الطلب:

أ- سنفترض أنَّ A هو حادث حصولنا على كرتين من لونٍ واحدٍ، فعندئذٍ يكون لدينا:

$$A = \{(r_1,r_2),(r_2,r_1),(b_1,b_2),(b_2,b_1)\}$$

ب- سنفترض أنَّ B هو حادث حصولنا على أرقام مختلفة، فعندئذٍ يكون لدينا:

$$B=\{(r_1,r_2),(r_2,r_1),(b_1,b_2),(b_2,b_1),(r_1,b_2),(r_2,b_1),(b_1,r_2),(b_2,r_1)\}$$

ج- سنفترض أنَّ C هو حادث حصولنا على مجموع للأرقام يساوي 4، فعندئذٍ يكون لدينا:

$$C=\{(r_2,b_2),\ (b_2,r_2)\}$$

د- سنفترض أنَّ D هو حادث حصولنا على مجموع للأرقام يساوي 3، فعندئذٍ يكون لدينا:

$$D=\{(r_1,r_2),(r_2,r_1),(b_1,b_2),(b_2,b_1),(r_1,b_2),(r_2,b_1),(b_1,r_2),(b_2,r_1)\}$$

هـ- سنفترض أنَّ E هو حادث حصولنا على مجموع للأرقام أصغر من 5، فعندئذٍ يكون لدينا:

$$E=\left\{\begin{matrix}(r_1,r_2),(r_2,r_1),(b_1,b_2),(b_2,b_1),(r_1,b_1),(r_1,b_2),\\(r_2,b_1),(r_2,b_2),(b_1,r_1),(b_1,r_2),(b_2,r_1),(b_2,r_2)\end{matrix}\right\}=\Omega$$

و- سنفترض أنَّ F هو حادث حصولنا على مجموع للأرقام أصغر من 2، فعندئذٍ نجد أنَّ:

$$F=\{\}=\varnothing$$

ز- سنفترض أنَّ G هو حادث حصولنا على كرات تحمل أرقاماً فرديةً، فعندئذٍ يكون لدينا:

$$G=\{(r_1,b_1),\ (b_1,r_1)\}$$

٥ - ٥
الدالَّة الاحتمالية ومبدأ لابلاس في الحساب الاحتمالي

لقد لاحظنا سابقاً أنَّه لدى تجربة عشوائية ما ستنشأ لدينا حوادث، وفي هذا الإطار قد يسأل المرء عن احتمال تحقّق حادث معيَّن متعلِّق بهذه التجربة، فعلى سبيل المثال في تجربة إلقاء حجر النرد لمرَّة واحدة فقط قد يُطرح السؤال الآتي:

ما هو احتمال حصولنا على عدد فردي في تجربة إلقاء حجر النرد لمرَّة واحدة فقط؟

يبدو للوهلة الأولى أنَّ الإجابة على هذا السؤال بسيطة جداً، ولكن في الواقع الأمر ليس كذلك لأسباب عديدة، منها على سبيل المثال:

١- كيف يمكننا تعيين احتمال الحادث الابتدائي في تجربة عشوائية ما؟

٢- ما هي الأداة التي ستقوم بحساب هذا الاحتمال؟

٣- ما هي الخصائص التي يجب على هذه الأداة تحقيقها حتى تصبح عديمة التناقض عند التعميم؟

إذاً علينا الإجابة على الأسئلة الثلاث السابقة قبل البدء في الإجابة على السؤال المطروح أعلاه، ولكن قبل البدء بتقديم هذه الإجابات على الأسئلة الثلاث السابقة نودّ التنويه إلى الملاحظة الآتية.

٥-٥-١- ملاحظة

عندما نذكر مستقبلاً أنَّ الأدوات المستخدمة في التجربة (مثل: قطع نقود أو حجارة نرد أو كرات أو بطاقات أو ...) متوازنة، فإنَّنا نقصد بذلك أنَّ هذه الأدوات مصنَّعة من مادة متجانسة الكثافة بحيث يصبح لكل نتيجة من نتائج التجربة العشوائية النصيب نفسه في الظهور.

الآن من أجل الإجابة على السؤال الأول يمكننا الجزم بأنَّ احتمال الحادث الابتدائي في تجربة عشوائية ما ليس من صلب عمل الاحتمالات، وذلك لأنَّ احتمال حادث ابتدائي متعلِّق بطبيعة المادة التي تجرى عليها التجربة، فعلى سبيل المثال:

أ- لو أخذنا تجربة قذف قطعة نقود متوازنة لمرَّة واحدة فقط، فعندئذٍ سيكون لكل نتيجة من نتائج التجربة العشوائية النصيب نفسه في الظهور، ومن ثمَّ احتمال كل حادث ابتدائي متعلِّق بهذه التجربة سوف يساوي $\frac{1}{2}$ لأنَّه لدينا نتيجتين فقط.

لكن في حال ليس لدينا معلومات حول تجانس المادة التي صُنِّعت منها قطعة النقود، فعندئذٍ **لا يمكننا** الادعاء أنَّ احتمال كل حادث ابتدائي متعلِّق بهذه التجربة يساوي $\frac{1}{2}$، فقد يكون مركز ثقل القطعة منحازاً إلى أحد وجهي القطعة بسبب عدم تجانس المادة المكوِّنة للقطعة، ومنه يصبح للوجه الآخر احتمال أكبر في الظهور للأعلى، وهكذا فإذا لم نُعطِ الاحتمال لظهور كل وجهٍ من الوجهين سوف لن يكون بإمكاننا حساب احتمالات لحوادث متعلِّقة بهذه التجربة العشوائية.

ب- لو أخذنا تجربة إلقاء حجر نرد **متوازن** لمرَّة واحدة فقط، فعندئذٍ سيكون احتمال كل حادث ابتدائي متعلِّق بهذه التجربة يساوي $\frac{1}{6}$ لأنَّه لدينا ست نتائج، ولكن في حال أنَّه ليس لدينا معلومات حول تجانس المادة التي صنِّع منها حجر النرد فإنَّه **لا يمكننا** الادعاء أنَّ احتمال كل حادث ابتدائي متعلِّق بهذه التجربة يساوي $\frac{1}{6}$، فقد يكون مركز ثقل حجر النرد ليس في مركزه بسبب عدم تجانس المادة المكوِّنة لحجر النرد، ومن ثمَّ يصبح لكل وجه من الوجوه نصيب مختلف في الظهور عن الآخر، وما لم يُعطى احتمال الظهور لكل وجهٍ من الوجوه الستة سوف لن يكون بإمكاننا حساب احتمالات لحوادث متعلِّقة بهذه التجربة العشوائية.

أمَّـا للإجابة على السؤالين الثاني والثالث فقد اقترح تقديم دالَّة تقوم بذلك على أن تحقِّق شروطاً محدّدة، وفي هذا الإطار بُذِلَت من أجلهما محاولات جادة من قبل بعض علماء الرياضيات (**وعلماء الاحتمالات على وجه الخصوص**)، وقد تراوحت نتائجهم ما بين عدم الدِقّة حيناً والتخصيص حيناً آخر، إلى أن جاء التعميم في النصف الأول من القرن العشرين (**وذلك في عام 1933**) على يد الرياضياتي الروسي كلموغوراف، ولكن في كتابنا هذا سوف لن نتطرق إلى هذه الدراسات وسنكتفي بتقديم نموذج بسيط يتوافق مع التخصيص الذي فرضناه على فضاء الحوادث الابتدائية Ω (**أي عندما تكون مجموعة نتائج التجربة منتهية**).

ليكن لدينا تجربة عشوائية بفضاء حوادث ابتدائية $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ منتهٍ مع n من $\mathbb{N}$ مُثبّت، وعلاوةً على ذلك سنفترض أنَّ لجميع نتائج التجربة العشوائية النصيب نفسه في الظهور، وهذه الصفة تنتج من مواقف عمليّة كثيرة إذ إنَّه يمكن الاستفادة من خاصية التجانس للمادة التي تُجرى عليها التجربة لتحقيق هذه الخاصيّة. أخيراً سنأخذ دالَّة حقيقية $\mathbf{P}$ (**أي مجالها المقابل** $\mathbb{R}$) معرَّفة على 2^Ω، أي أنَّ:

$$\mathbf{P} : 2^\Omega \longrightarrow \mathbb{R} \; ; \; A \mapsto \mathbf{P}(A)$$

ومُحقّقة لما يلي:

١- لدينا $\mathbf{P}(\varnothing) = 0$.

٢- من أجل أية حوادث A_1 و A_2 و ... و A_k من 2^Ω متنافية مثنى مثنى مع $2 \le k$ عدد طبيعي مثبّت لدينا:

$$\mathbf{P}\left(\bigcup_{i=1}^{k} A_i\right) = \sum_{i=1}^{k} \mathbf{P}(A_i)$$

٣- لدينا $\mathbf{P}(\Omega) = 1$.

إنَّ الدالَّة P المُحقِّقة للشروط الثلاثة السابقة تُدعى **دالَّة احتمالية**.

٥-٥-٢- ملاحظات:

١- من الشرطين (١) و (٣) السابقين نلاحظ أنَّ مدى الدالَّة P (مجموعة قيمها) هي الفترة $[0\,,1]$ فقط، ومن ثمَّ يكون من أجل أي حادث A من 2^{Ω} لدينا العلاقة الآتية محقَّقة:

$$0 \le \mathbf{P}(A) \le 1 \qquad [5\text{-}4]$$

٢- لو أخذنا الحوادث البسيطة $\{\omega_1\}$ و $\{\omega_2\}$ و... و $\{\omega_n\}$، فعندئذٍ سيكون لجميع هذه الحوادث الاحتمال نفسه (لأنَّ لجميع نتائج التجربة العشوائية النصيب نفسه في الظهور)، أي أنَّ:

$$\mathbf{P}(\{\omega_1\}) = \mathbf{P}(\{\omega_2\}) = \ldots = \mathbf{P}(\{\omega_n\})$$

وبالتالي بحسب الشرط الثاني والثالث ينتج لدينا الآتي:

$$1 = \mathbf{P}(\Omega) = P\left(\bigcup_{i=1}^{n}\{\omega_i\}\right) = \sum_{i=1}^{n}\mathbf{P}(\{\omega_i\})$$

$$= \underbrace{\mathbf{P}(\{\omega_1\}) + \mathbf{P}(\{\omega_1\}) + \ldots + \mathbf{P}(\{\omega_1\})}_{n-times} = n\,\mathbf{P}(\{\omega_1\})$$

والتي ينتج عنها أنَّ $\mathbf{P}(\{\omega_1\}) = \dfrac{1}{n}$، وبالتالي يكون لدينا:

$$\mathbf{P}(\{\omega_1\}) = \mathbf{P}(\{\omega_2\}) = \ldots = \mathbf{P}(\{\omega_n\}) = \frac{1}{n} \qquad [5\text{-}5]$$

ومن ثمَّ احتمال أي حادث A متعلِّق بالتجربة العشوائية سيكون مساوياً إلى النسبة $\dfrac{|A|}{|\Omega|}$، وهذا يعني أنَّه يمكننا أن نكتب:

$$\mathbf{P}(A) = \frac{|A|}{|\Omega|} \quad ;\ \forall A \in 2^{\Omega} \qquad [5\text{-}6]$$

وهذه العلاقة صاغها لابلاس على النحو الآتي أيضاً:

احتمال الحادث الذي قيد الدراسة = عدد الحالات الملائمة للحادث قيد الدراسة / عدد الحالات الممكنة للتجربة

إنَّ الدالَّة الاحتمالية P المُعطاة بالعلاقة الأخيرة (والتي وضعها لابلاس) يُطلَق عليها اسم **مبدأ لابلاس من أجل الحالات المتساوية الإمكانية**، وعُرِفَت فيما بَعد باسم "**التعريف التقليدي للاحتمال**" Classical Definition of Probability.

إنَّ التعامل مع هذا التعريف للاحتمال سيقود المرء في كثير من الحالات إلى استخدام التحليل التوافقي (استخدام التراتيب والتوافيق) لحل المسائل، ومن ثمَّ سنلاحظ أنَّ التحليل التوافقي سيقوم بدورٍ مهم في الحساب الاحتمالي لمسائل تتوافق وحالتنا هذه. هذا من جانب، ومن جانب آخر يجب على المرء ألاّ يتجاوز فرضيات هذا التعريف إذ إنَّه يصبح عديم الفائدة عندما يكون فضاء الحوادث الابتدائية غير منتهٍ أو عندما يكون للحوادث الابتدائية احتمالات مختلفة.

◄ ٥-٥-٣- أمثلة

١- بالعودة إلى المثال /١/ من (٥-٤-٤) حيث لدينا تجربة رمي حجر نرد لمرّة واحدة فقط، وسنفترض علاوةً على ما سبق أنَّ حجر النرد متوازن، فعندئذٍ يكون استخدام مبدأ لابلاس من أجل حساب احتمالات لحوادث متعلِّقة بهذه التجربة قابلاً للتطبيق (لأنَّ عدد نتائج هذه التجربة منتهٍ ولجميع الأوجه النصيب نفسه في الظهور)، ولذلك سنقوم بحساب احتمالات الحوادث التي تمَّ تعيينها في ذلك المثال، حيث لدينا:

أ- الحادث $A=\{1,2\}$ (والذي يعبرِّ عن حصولنا على عدد أصغر من 3)، ومن ثمَّ يكون احتمال هذا الحادث يساوي:

$$\mathbf{P}(A)=\frac{|A|}{|\Omega|}=\frac{2}{6}=0.\bar{3}$$

ب- الحادث $B=\{5,6\}$ (والذي يعبرِّ عن حصولنا على عدد أكبر من 4)، ومن ثمَّ يكون احتمال هذا الحادث يساوي:

$$\mathbf{P}(B)=\frac{|B|}{|\Omega|}=\frac{2}{6}=0.\bar{3}$$

ج- الحادث $C=\{3,4\}$ (والذي يعبرِّ عن حصولنا على عدد أكبر أو يساوي 3 وأصغر أو يساوي 4)، ومن ثمَّ يكون احتمال هذ الحادث يساوي:

$$\mathbf{P}(C)=\frac{|C|}{|\Omega|}=\frac{2}{6}=0.\bar{3}$$

د- الحادث $D=\varnothing$ (والذي يعبرِّ عن حصولنا على عدد أصغر من 1)، ومن ثمَّ يكون احتمال هذ الحادث يساوي:

$$\mathbf{P}(D)=\frac{|D|}{|\Omega|}=\frac{0}{6}=0$$

هـ- الحادث $E=\Omega$ (والذي يعبرِّ عن حصولنا على عدد أكبر أو يساوي 1)، ومن ثمَّ يكون احتمال هذ الحادث يساوي:

$$\mathbf{P}(E)=\frac{|E|}{|\Omega|}=\frac{|\Omega|}{|\Omega|}=\frac{6}{6}=1$$

و- الحادث $F=\{1,3,5\}$ (والذي يعبِّر عن حصولنا على عدد فردي)، ومن ثمَّ يكون احتمال هذ الحادث يساوي:

$$\mathbf{P}(F)=\frac{|F|}{|\Omega|}=\frac{3}{6}=\frac{1}{2}=0.5$$

٢- بالعودة إلى المثال /٢/ من (٥-٤-٤) السابق حيث لدينا تجربة سحب كرتين على التتالي دون إرجاع إلى الصندوق، وفي هذه المسألة نلاحظ أنَّ استخدام مبدأ لابلاس من أجل حساب احتمالات لحوادث متعلِّقة بالتجربة قابلاً للتطبيق دون إضافة أية شروط إضافية (لأنَّ عدد نتائج هذه التجربة منتهٍ ولجميع الكرات النصيب نفسه في الظهور)، ولذلك سنقوم بحساب احتمالات الحوادث التي تمَّ تعيينها في ذلك المثال، حيث لدينا:

أ- الحادث $A=\{(r_1,r_2),(r_2,r_1),(b_1,b_2),(b_2,b_1)\}$ يعبِّر عن حصولنا على كرتين من لون واحد، واحتماله يساوي:

$$\mathbf{P}(A)=\frac{|A|}{|\Omega|}=\frac{4}{12}=\frac{1}{3}=0.\overline{3}$$

ب- الحادث:

$$B=\{(r_1,r_2),(r_2,r_1),(b_1,b_2),(b_2,b_1),(r_1,b_2),(r_2,b_1),(b_1,r_2),(b_2,r_1)\}$$

يعبِّر عن حصولنا على أرقام مختلفة، واحتماله يساوي:

$$\mathbf{P}(B)=\frac{|B|}{|\Omega|}=\frac{8}{12}=\frac{2}{3}=0.\overline{6}$$

ج- الحادث $C=\{(r_2,b_2),\ (b_2,r_2)\}$ يعبِّر عن حصولنا على مجموع للأرقام يساوي 4، واحتماله يساوي:

$$\mathbf{P}(C)=\frac{|C|}{|\Omega|}=\frac{2}{12}=\frac{1}{6}=0.1\overline{6}$$

د- الحادث:

$$D=\{(r_1,r_2),(r_2,r_1),(b_1,b_2),(b_2,b_1),(r_1,b_1),(r_1,b_2),(r_2,b_1),(b_1,r_1),(b_1,r_2),(b_2,r_1)\}$$

يعبِّر عن حصولنا على مجموع للأرقام يساوي 3، واحتماله يساوي:

$$\mathbf{P}(D)=\frac{|D|}{|\Omega|}=\frac{10}{12}=\frac{5}{6}=0.8\overline{3}$$

هـ- الحادث $E=\Omega$ يعبِّر عن حصولنا على مجموع للأرقام أصغر من 5 واحتماله يساوي:

$$\mathbf{P}(E)=\frac{|E|}{|\Omega|}=\frac{12}{12}=1$$

و- الحادث $F = \varnothing$ يعبرِّ عن حصولنا على مجموع للأرقام أصغر من 2 واحتماله يساوي:

$$\mathbf{P}(F) = \frac{|F|}{|\Omega|} = \frac{0}{12} = 0$$

ز- الحادث $G = \{(r_1, b_1), (b_1, r_1)\}$ يعبرِّ عن حصولنا على كرات تحمل أرقاماً فردية واحتماله يساوي:

$$\mathbf{P}(G) = \frac{|G|}{|\Omega|} = \frac{2}{12} = \frac{1}{6} = 0.1\overline{6}$$

٣- في مستوصف يوجد كادر طبي مكوَّن من طبيبين وثلاث ممرضات. قامت إدارة المستوصف بتشكيل لجنة مكوَّنة من ثلاثة أشخاص اختيروا عشوائياً من هذا الكادر، فإذا علمت أنَّ لكل شخص في هذا الكادر النصيب نفسه في الاختيار، فما هو احتمال وجود الطبيبين في هذه اللجنة؟

الحل: لنرمز بـ A لحادث وجود طبيبين في هذه اللجنة. عندئذٍ بسبب أنَّه لم يذكر شيء عن كيفية سحب عناصر هذه اللجنة من الكادر الطبي، أهو شخص بعد الآخر أم دفعة واحدة، فإنَّه علينا مناقشة الحالتين الآتيتين:

أ- إذا كان السحب قد تمَّ لشخص بعد الآخر، فعندئذٍ يكون عدد الحالات المكنة للتجربة هو:

$$|\Omega| = 5P3 = 5 \times 4 \times 3 = 60$$

وأمَّا عدد الحالات الملائمة للحادث A يساوي:

$$|A| = (2P2) \cdot (3P1) = 2 \times 3 = 6$$

ومن ثمَّ يكون الاحتمال المطلوب يساوي:

$$\mathbf{P}(A) = \frac{|A|}{|\Omega|} = \frac{6}{60} = 0.1$$

ب- إذا كان السحب تمَّ على دفعة واحدة، فعندئذٍ يكون عدد الحالات المكنة للتجربة هو:

$$|\Omega| = 5C3 = 10$$

وأمَّا عدد الحالات الملائمة للحادث A يساوي:

$$|A| = (2C2) \cdot (3C1) = 1 \times 3 = 3$$

ومن ثمَّ يكون الاحتمال المطلوب يساوي:

$$\mathbf{P}(A) = \frac{|A|}{|\Omega|} = \frac{3}{10} = 0.3$$

٤- يوجد في صندوقين A وB كرات **متماثلة تماماً** وبألوانٍ مختلفةٍ، وبحيث يحوي الصندوق A كرة بيضاء وأخرى سوداء فقط، وأمّا الصندوق B فإنَّه يحوي كرة بيضاء وكرة سوداء وأخرى حمراء. نقوم بخلط الكرات في الصندوق A جيداً ومن ثمَّ نسحب عشوائياً كرة من هذا الصندوق ونضعها في الصندوق الثاني دون رؤيتها، وبعد ذلك نقوم بخلط كرات الصندوق B جيداً ومن ثمَّ نسحب عشوائياً كرة منه، فما هو احتمال أن تكون الكرة المسحوبة أخيراً بيضاء؟

الحل: بما أنَّ جميع الكرات متماثلة فإنَّ ذلك يعني أنَّ لجميع الحوادث الابتدائية الناتجة عن هذه التجربة النصيب نفسه في الظهور، ومن ثمَّ سيكون لجميع الحوادث الابتدائية الاحتمال نفسه.

الآن لتعيين فضاء الحوادث الابتدائية سنرمز للكرة البيضاء والسوداء في الصندوق A بـ w_A و b_A على الترتيب، في حين سنرمز للكرة البيضاء والسوداء والحمراء في الصندوق B بـ w_B و b_B و r_B على الترتيب. عندئذٍ يكون لفضاء الحوادث الابتدائية العرض الآتي:

$$\Omega = \left\{(w_A, w_A), (w_A, w_B), (w_A, b_B), (w_A, r_B), (b_A, b_A), (b_A, w_B), (b_A, b_B), (b_A, r_B)\right\}$$

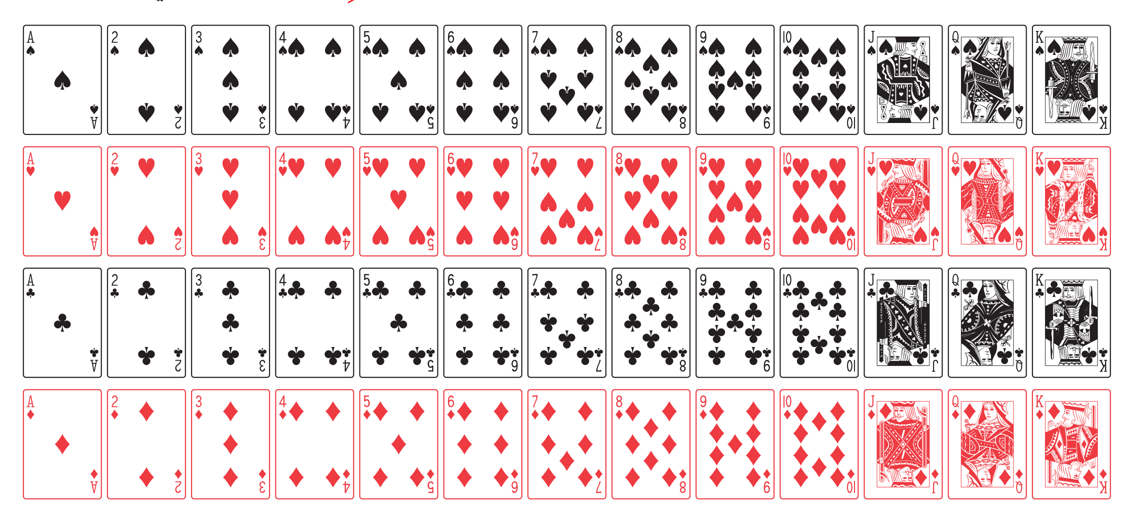
حيث ترمز المركبة الأولى من الثنائية إلى نتيجة السحب الأول بينما ترمز المركبة الثانية من الثنائية إلى نتيجة السحب الثاني. نلاحظ هنا أنَّ فضاء الحوادث الابتدائية لهذه التجربة منته، ومن ثمَّ يمكننا تطبيق مبدأ لابلاس في الاحتمالات لحساب الاحتمال المطلوب، فلو افترضنا أنَّ C هو حادث سحب كرة بيضاء من الصندوق الثاني بَعد إضافة كرة إليه من الصندوق الأول فإنَّه يمكننا أن نكتب:

$$C = \left\{(w_A, w_A), (w_A, w_B), (b_A, w_B)\right\}$$

ومن ثمَّ يكون الاحتمال المطلوب يساوي:

$$\mathbf{P}(C) = \frac{|C|}{|\Omega|} = \frac{3}{8} = 0.375$$

٥- ليكن لدينا بطاقات لعب Playing Carts مكوَّنة من 52 بطاقةٍ لها العرض الآتي:



علماً أنَّه لدينا أربعة أنواع في هذا النوع من البطاقات هي: ♦، ♥، ♠ و ♣ وتُدعى دينار Diamond، قلب Heart ، بستوني Spade و زهر Club على الترتيب.

والآن. إذا علمت أنَّ لجميع البطاقات النصيب نفسه في الاختيار (أو السحب)، وقمنا بسحب عشوائي لأربع بطاقات دفعةً واحدةً من هذه المجموعة، فما هو احتمال أن يكون:

أ- لدينا بطاقة سوداء واحدة فقط؟

ب- لدينا ثلاث صور؟

ج- جميع البطاقات المسحوب آسات Aces (والممثَّلة بنقطة واحدة)؟

د- لدينا من كل نوع بطاقة؟

الحل: بما أنَّ فضاء الحوادث الابتدائية لهذه التجربة منتهٍ ولجميع البطاقات النصيب نفسه في الاختيار فإنَّه يمكننا تطبيق مبدأ لابلاس في حساب الاحتمالات المطلوبة، حيث نلاحظ في هذه المسألة أنَّ عدد الحالات المكنة للتجربة يساوي إلى عدد الطرائق التي يمكننا فيها اختيار 4 بطاقات من 52 بطاقة، وبما أنَّ الترتيب هنا ليس له أهمية فإنَّ عدد هذه الطرائق يساوي:

$$52C4 = \frac{52!}{4! \cdot (52-4)!} = 270725$$

والآن من أجل الإجابة على الطلب:

أ- سنرمز بـ A لحادث حصولنا على بطاقة سوداء واحدة فقط، فعندئذٍ يكون لدينا عدد الحالات الملائمة لهذا الحادث يساوي إلى عدد الطرائق التي يمكننا فيها اختيار بطاقة واحدة من 26 بطاقة سوداء واختيار 3 بطاقات من 26 بطاقة حمراء (حيث الترتيب هنا ليس له أهمية)، وعدد هذه الطرائق يساوي:

$$(26C1)\cdot(26C3) = 26 \times 2600 = 67600$$

ومن ثمَّ الاحتمال المطلوب يساوي:

$$\mathbf{P}(A) = \frac{(26C1)\cdot(26C3)}{52C4} = \frac{67600}{270725} = 0.25$$

ب- سنرمز بـ B لحادث حصولنا على ثلاث صور، فعندئذٍ يكون لدينا عدد الحالات الملائمة لهذا الحادث يساوي إلى عدد الطرائق التي يمكننا فيها اختيار ثلاث بطاقات من 12 صورة، واختيار البطاقة الرابعة من الـ 40 بطاقة المتبقية (حيث الترتيب هنا ليس له أهمية)، وعدد هذه الطرائق يساوي:

$$(12C3)\cdot(40C1) = 440 \times 40 = 17600$$

ومن ثمَّ الاحتمال المطلوب يساوي:

$$\mathbf{P}(B) = \frac{(12C3)\cdot(40C1)}{52C4} = \frac{17600}{270725} = 0.065$$

ج- سنرمز بـ C لحادث حصولنا على أربعة آسات، فعندئذٍ يكون لدينا عدد الحالات الملائمة لهذا الحادث يساوي إلى عدد الطرائق التي يمكننا فيها اختيار أربع بطاقات من 4 فقط (حيث الترتيب هنا ليس له أهمية)، وعدد هذه الطرائق يساوي $4C4 = 1$، ومن ثمَّ الاحتمال المطلوب يساوي:

$$\mathbf{P}(C) = \frac{4C4}{52C4} = \frac{1}{270725} = 0.0000037$$

د- سنرمز بـ D لحادث حصولنا على بطاقة من كل نوع، فعندئذٍ يكون لدينا عدد الحالات الملائمة لهذا الحادث يساوي إلى عدد الطرائق التي يمكننا فيها اختيار بطاقة واحدة من 13 بطاقة وبالمثل بالنسبة لبقية البطاقات الأخرى (حيث الترتيب هنا ليس له أهمية)، وعدد هذه الطرائق يساوي:

$$(13C1)\cdot(13C1)\cdot(13C1)\cdot(13C1)=13\times 13\times 13\times 13=28561$$

ومن ثمَّ الاحتمال المطلوب يساوي:

$$\mathbf{P}(D)=\frac{(13C1)\cdot(13C1)\cdot(13C1)\cdot(13C1)}{52C4}=\frac{28561}{270725}=0.1055$$

٥-٥-٣- بعض خصائص الدالَّة الاحتمالية

إنَّ المقولات الآتية (سنقدِّمها دون برهان) تعرض لنا بعض الخصائص البسيطة للدالَّة الاحتمالية.

١- من أجل أي حادث A من 2^{Ω} لدينا:

$$\mathbf{P}(\bar{A})=1-\mathbf{P}(A) \qquad [5\text{-}7]$$

٢- من أجل أي حادثين A وB من 2^{Ω} لدينا:

$$\mathbf{P}(B\setminus A)=\mathbf{P}(B)-\mathbf{P}(A\cap B) \qquad [5\text{-}8]$$

٣- من أجل أي حادثين A وB من 2^{Ω} مع $A\subset B$ (أي أنَّ تحقّق B يقتضي تحقّق A حتماً) سيكون لدينا:

$$\mathbf{P}(B)\geq \mathbf{P}(A)$$

وهذه الخاصية تُعرَف باسم **خاصية الاطراد** Monotone Property للدالَّة الاحتمالية، وعلاوةً على ذلك يكون لدينا في هذه الحالة:

$$\mathbf{P}(B\setminus A)=\mathbf{P}(B)-\mathbf{P}(A) \qquad [5\text{-}9]$$

٤- من أجل أية ثلاثة حوادث A وB وC من 2^{Ω} لدينا:

$$\begin{aligned}\mathbf{P}(A\cup B\cup C)=\mathbf{P}(A)+\mathbf{P}(B)+\mathbf{P}(C)\\ -\mathbf{P}(A\cap B)-\mathbf{P}(A\cap C)-\mathbf{P}(B\cap C)\\ +\mathbf{P}(A\cap B\cap C)\end{aligned} \qquad [5\text{-}10]$$

والذي يُدعى **قانون الجمع** Addition Law في الاحتمالات، وفي الحالة الخاصّة إذا كانت الحوادث A وB وC متنافية مثنى مثنى، فعندئذٍ يصبح لدينا:

$$\mathbf{P}(A\cup B\cup C)=\mathbf{P}(A)+\mathbf{P}(B)+\mathbf{P}(C)$$

وفي حال كان لدينا حادثين فقط A وB من 2^{Ω}، فعندئذٍ يكون لدينا:

$$\mathbf{P}(A\cup B)=\mathbf{P}(A)+\mathbf{P}(B)-\mathbf{P}(A\cup B) \qquad [5\text{-}11]$$

وعندما يكون الحادثان A وB متنافيين (أي أنَّ $A \cap B = \varnothing$) فعندئذٍ يصبح لدينا:

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$$

◄ ٥-٥-٤- أمثلة

١- ليكن لدينا A وB وC حوادث من فضاء حوادث ابتدائية Ω، ولنفترض أنَّ:

$$\mathbf{P}(A \setminus B) = 0.25 \quad \& \quad \mathbf{P}(B \setminus A) = 0.30 \quad \& \quad \mathbf{P}(C \setminus A) = 0.10$$

$$\mathbf{P}(A \cap B) = 0.15 \quad \& \quad \mathbf{P}(A \cap C) = 0.10 \quad \& \quad \mathbf{P}(B \cap C) = 0.15$$

$$\mathbf{P}(A \cap B \cap C) = 0.05$$

ولنقم بحساب الاحتمالين $\mathbf{P}(A \cup B)$ و $\mathbf{P}(A \cup B \cup C)$.

الحل: نعلم أنَّ $\mathbf{P}(A \cup B \cup C)$ يحسب بالعلاقة [5-10]، وكذلك $\mathbf{P}(A \cup B)$ يحسب باستخدام العلاقة [5-11]، ولذلك علينا أولاً حساب الاحتمالات $\mathbf{P}(A)$، $\mathbf{P}(B)$ و $\mathbf{P}(C)$ حيث لدينا من العلاقة [5-8] ما يلي:

$$\mathbf{P}(A) = \mathbf{P}(A \setminus B) + \mathbf{P}(A \cap B) = 0.25 + 0.15 = 0.40$$

$$\mathbf{P}(B) = \mathbf{P}(B \setminus A) + \mathbf{P}(A \cap B) = 0.30 + 0.15 = 0.45$$

$$\mathbf{P}(C) = \mathbf{P}(C \setminus A) + \mathbf{P}(A \cap C) = 0.10 + 0.10 = 0.20$$

ومن ثمَّ يكون لدينا:

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B) = 0.40 + 0.45 - 0.15 = 0.70$$

وكذلك نجد أنَّ:

$$\begin{aligned}\mathbf{P}(A \cup B \cup C) &= \mathbf{P}(A) + \mathbf{P}(B) + \mathbf{P}(C) \\ &\quad - \mathbf{P}(A \cap B) - \mathbf{P}(A \cap C) - \mathbf{P}(B \cap C) + \mathbf{P}(A \cap B \cap C) \\ &= 0.40 + 0.45 + 0.20 - 0.15 - 0.10 - 0.15 + 0.05 = 0.70\end{aligned}$$

٢- لنأخذ تجربة إلقاء قطعة نقود متوازنة لثلاث مرَّات متتالية، ولنقم بتعيين فضاء الحوادث الابتدائية لهذه التجربة العشوائية، ومن ثمَّ حساب احتمال الحصول على صورة واحدة على الأقل في هذه التجربة.

الحل: إنَّ مجموعة نتائج هذه التجربة العشوائية هي:

$$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

وأمَّا لحساب الاحتمال المطلوب فإنَّنا سنفترض أنَّ A هو حادث الحصول على صورة واحدة على الأقل في هذه التجربة، فعندئذٍ يكون لدينا:

$$A = \{HHH, HHT, HTH, THH, HTT, THT, TTH\}$$

وبما أنَّ قطعة النقود متوازنة فإنه سيكون لجميع النتائج النصيب نفسه في الظهور، ومن ثمَّ يصبح لدينا بحسب مبدأ لابلاس في الاحتمالات ما يلي:

$$\mathbf{P}(A) = \frac{|A|}{|\Omega|} = \frac{7}{8} = 0.875$$

لاحظ هنا كان بإمكاننا حسابه باستخدام حساب احتمال الحادث المتمم حيث لدينا:

$$\bar{A} = \{TTT\} \quad \Rightarrow \quad \mathbf{P}(\bar{A}) = \frac{|\bar{A}|}{|\Omega|} = \frac{1}{8} = 0.125$$

ومن ثمَّ يكون لدينا بحسب العلاقة [5-7] ما يلي:

$$\mathbf{P}(A) = 1 - \mathbf{P}(\bar{A}) = 1 - 0.125 = 0.875$$

٥ - ٦
الاحتمالات الشرطية

لقد قمنا فيما سبق بالتعرّف على مفهوم الحادث وكيفية حساب احتماله بوساطة الدالّة الاحتمالية، ولكن قد يصادفنا في كثير من الحالات حساب احتمالات لحوادث ذات طبيعة شرطية، فعلى سبيل المثال:

أ- لدى رمي حجر نرد متوازن، فما هو احتمال حصولنا على العدد 1 علماً أنَّنا قد حصلنا على عدد فردي؟

ب- لدى قذف قطعة نقود متوازنة لثلاث مرّات متتالية، فما هو احتمال حصولنا على صورة في المرّة الثانية علماً أنَّنا قد حصلنا على صورتين تماماً؟

نلاحظ هنا أنَّ المطلوب حساب احتمال حادث إذا عُلِمَ تَحقُّق وقوع حادث آخر سابق له، ولتوضيح ذلك أكثر لنأخذ تجربة إلقاء حجر نرد متوازن لمرَّة واحدة فقط، ولنطرح السؤال الآتي:

ما هو احتمال الحصول على الرقم 1 علماً أنَّنا قد حصلنا على عدد فردي؟

من أجل الإجابة على هذا السؤال سنرمز بـ A لحادث الحصول على العدد 1، وأمّا حادث الحصول على عدد فردي فسنرمز له بـ B. عندئذٍ يكون لـدينا:

$$A = \{1\} \qquad \& \qquad B = \{1,3,5\}$$

فنجد أنَّ $A \cap B$ هو الحادث الذي يعبِّر عن حصولنا على العدد 1 وعدد فردي، ومن جهةٍ أخرى نلاحظ أنَّه بعد أن علمنا أنَّنا قد حصلنا على عدد فردي فإنَّ فضاء الحوادث الابتدائية Ω قد اخْتُزِلَ إلى الحادث B لأنَّ الحالات المكنة للتجربة ستصبح 1 و3 و5 فقط، ومن ثمَّ بحسب مبدأ لابلاس في الاحتمالات فإنَّ الاحتمال المطلوب سيكون مساوياً النسبة الآتية:

$$\frac{\text{عدد الحالات الملائمة للحادث}}{\text{عدد الحالات المكنة للتجربة}} = \frac{|A \cap B|}{|B|} = \frac{1}{3}$$

فلو قمنا الآن بتقسيم البسط والمقام على $|\Omega|$ فإنَّه يصبح لدينا:

$$\frac{\frac{1}{6}}{\frac{3}{6}} = \frac{\frac{|A \cap B|}{|\Omega|}}{\frac{|B|}{|\Omega|}} = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}$$

إذاً احتمال الحصول على العدد 1 علماً أنّنا قد حصلنا على عدد فردي يساوي قيمة النسبة $\frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}$ ويساوي $\frac{1}{3}$. بالطبع يُشترط هنا أن يكون $0 < \mathbf{P}(B)$ لأنَّ النسبة $\frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}$ لا معنى لها عندما يكون $\mathbf{P}(B) = 0$.

٥-٦-١ - تعريف (الاحتمال الشرطي لحادث)

Conditional Probability of an Event

ليكن Ω فضاء حوادث ابتدائية منته، ولنأخذ A وB حادثين من 2^Ω مع $0 < \mathbf{P}(B)$. عندئذٍ يُعرَّف **الاحتمال الشرطي** للحادث A علماً أنَّ الحادث B قد تحقَّق وقوعه (ويرمز له $\mathbf{P}(A \mid B)$) على أنَّه قيمة النسبة $\frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}$، أي أنَّ:

$$\mathbf{P}(A \mid B) := \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} \qquad [5-12]$$

٥-٦-١-١- ملاحظات

١- في الحالة الخاصة عندما يكون لجميع نتائج التجربة العشوائية النصيب نفسه في الظهور، فإنَّه يمكننا حساب الاحتمال الشرطي السابق من خلال العلاقة الآتية:

$$\mathbf{P}(A \mid B) = \frac{|A \cap B|}{|B|} \quad ; A \in 2^\Omega, |B| > 0 \qquad [5-13]$$

٢- إذا كان $0 < \mathbf{P}(A)$ فإنَّه يمكننا أن نكتب أيضاً:

$$\mathbf{P}(B \mid A) = \frac{|A \cap B|}{|A|} \quad ; B \in 2^\Omega$$

٣- من أجل A حادث ما من 2^Ω مع $0 < \mathbf{P}(A)$، فإنّه من تعريف الاحتمال الشرطي سينتج لدينا تَحقّق البنود الآتية:

أ- لدينا $\mathbf{P}(\varnothing \mid A) = 0$ وكذلك $\mathbf{P}(\Omega \mid A) = 1$.

ب- من أجل أي حادثين متنافيين B وC من 2^Ω لدينا:

$$\mathbf{P}\left((B \cup C) \mid A\right) = \mathbf{P}(B \mid A) + \mathbf{P}(C \mid A)$$

ج- من أجل أي حادث B من 2^Ω لدينا:

$$\mathbf{P}\left(\bar{B} \mid A\right) = 1 - \mathbf{P}(B \mid A)$$

د- إذا كان B حادثاً ما من 2^{Ω} مع $0 < \mathbf{P}(B)$ فعندئذٍ يمكننا أن نكتب الآتي:

$$\mathbf{P}(A \cap B) = \mathbf{P}(A) \cdot \mathbf{P}(B \mid A) = \mathbf{P}(B) \cdot \mathbf{P}(A \mid B)$$

إنَّ العلاقة الأخيرة تُدعى **قانون الضرب في الاحتمالات**.

◄ ٥-٦-١-٣- مثال

لنأخذ تجربة إلقاء قطعة نقود متوازنة لثلاث مرَّات متتالية، ولنقم بحساب احتمال الحصول على صورة في الرمية الأخيرة علماً أنَّنا حصلنا على صورة واحدة على الأقل خلال هذه الرميات الثلاث.

الحل: من أجل ذلك سنفترض أنَّ A هو الحادث الذي يعبِّر عن حصولنا على صورة في الرمية الأخيرة، وB هو الحادث الذي يعبِّر عن حصولنا على صورة واحدة على الأقل خلال الرميات الثلاث، فعندئذٍ يكون لدينا:

$$A = \{HHH, HTH, THH, TTH\}$$

$$B = \{HHH, HHT, HTH, THH, TTH, THT, HTT\}$$

ومن مُعطيات المسألة لدينا قطعة النقود متوازنة، ومنه يكون الاحتمال الشرطي لـ A علماً أنَّ الحادث B قد تحقَّق يساوي:

$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} = \frac{4/8}{7/8} = \frac{4}{7}$$

►

من التطبيقات الهامّة للاحتمال الشرطي ما يُعرف باسم "**صيغة الاحتمال التام**" وكذلك "**صيغة بييز**" والتي سنقدّمهما تباعاً بعد التمهيد الآتي.

٥-٦-٢- تعريف (التجزئة للحادث الأكيد)

ليكن $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ فضاء حوادث ابتدائية منتهٍ، ولنأخذ Z_1 و Z_2 و ... و Z_k حوادث من 2^{Ω} علماً أنَّ $n \geq k$. عندئذٍ يُقال عن أسرة من الحوادث $\mathbf{Z} = \{Z_1, Z_2, \ldots, Z_k\}$ إنَّها تجزئة للحادث الأكيد Ω إذا كان ما يلي مُحقَّقاً:

أ- جميع الحوادث Z_1 و Z_2 و ... و Z_k ليست مستحيلة.

ب- الحوادث Z_1 و Z_2 و ... و Z_k متنافية مثنى مثنى.

ج- لدينا $\Omega = Z_1 \cup Z_2 \cup \ldots \cup Z_k$.

الشكل [5-4]

الشكل الجانبي [5-4] يعرض لنا تجزئة للحادث الأكيد Ω من أجل $k = 5$.

٥-٦-٣- صيغة الاحتمال التام وصيغة بييز

Total Probability Formula and Bays's Formula

ليكن $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ فضاء حوادث ابتدائية منته، ولنأخذ $\mathbf{Z} = \{Z_1, Z_2, \dots, Z_k\}$ مع $n \geq k$ تجزئة للحادث الأكيد Ω مع $0 < \mathbf{P}(Z_1)$ و $0 < \mathbf{P}(Z_2)$ و ... و $0 < \mathbf{P}(Z_k)$، فعندئذٍ:

أ- من أجل أي حادث A من 2^Ω يكون لدينا:

$$\mathbf{P}(A) = \sum_{i=1}^{k} \mathbf{P}(Z_i) \cdot \mathbf{P}(A \mid Z_i) \qquad [5\text{–}14]$$

وهذه العلاقة تـُعرف باسم "صيغة الاحتمال التام" Total Probability Formula.

ب- من أجل أي حادث A من 2^Ω مع $0 < \mathbf{P}(A)$، ومن أجل أية قيمة صحيحة $1 \leq j \leq k$ يكون لدينا:

$$\mathbf{P}(Z_j \mid A) = \frac{\mathbf{P}(Z_j) \cdot \mathbf{P}(A \mid Z_j)}{\sum_{i=1}^{k} \mathbf{P}(Z_i) \cdot \mathbf{P}(A \mid Z_i)} \qquad [5\text{–}15]$$

وهذه العلاقة الأخيرة تُعرف باسم " صيغة بييز في الاحتمالات" Bays's Formula، وتنسب إلى الإحصائي والفيلسوف الإنجليزي بييز Thomas Bayes (1701–1761).

◀ ٥-٦-٥- أمثلة

١- يوجد في عمادة السنة الأولى المشتركة ثلاثة مسارات لتدريس الطلبة هي: الإنساني، العلمي والصحي، فإذا علمنا أنَّ نسبة أعداد الطلبة في هذه المسارات هي % 30 و % 50 و % 20 على الترتيب، وأنَّ نسبة أعداد الطلبة المتميِّزين في كل من المسارات الثلاثة هو 0.10 و0.15 و0.20 على الترتيب، فإذا قمنا بسحب عشوائي لطالب من عمادة السنة الأولى المشتركة، فعندئذٍ:

أ- ما هو احتمال أن يكون الطالب الذي تمَّ اختياره متميّزاً؟

ب- ما هو احتمال أن يكون الطالب الذي تمَّ اختياره من المسار الإنساني إذا وجدنا أنَّه متميّز؟

✍ **الحل:** للإجابة على هذين السؤالين سنفترض أنَّ Z_1، Z_2 و Z_3 هو الحادث الذي يعبِّر عن كون الطالب الذي تمَّ اختياره من المسار الإنساني، العلمي والصحي على الترتيب، فنجد أنَّ هذه الحوادث الثلاثة تشكل تجزئة للحادث الأكيد Ω (الذي يمثِّل كل الطلبة في عمادة السنة الأولى المشتركة)، ومن ثمَّ بفرض A هو الحادث الذي يعبِّر عن كون الطالب الذي تمَّ اختياره من المتميِّزين، فإنَّه سيكون لدينا من أجل الطلب:

أ – احتمال أن يكون الطالب الذي تمَّ اختياره متميّزاً يساوي (بحسب صيغة الاحتمال التام):

$$\mathbf{P}(A)=\sum_{i=1}^{3}\mathbf{P}(Z_i)\cdot\mathbf{P}(A\mid Z_i)$$
$$=\frac{30}{100}\cdot\frac{10}{100}+\frac{50}{100}\cdot\frac{15}{100}+\frac{20}{100}\cdot\frac{20}{100}=\frac{1450}{10000}=0.145$$

ب- احتمال أن يكون الطالب الذي تمَّ اختياره من المسار الإنساني إذا وجد أنَّه متميّز (بحسب صيغة بييز في الاحتمالات):

$$\mathbf{P}(Z_1\mid A)=\frac{\mathbf{P}(Z_1)\cdot\mathbf{P}(A\mid Z_1)}{\sum_{i=1}^{3}\mathbf{P}(Z_i)\cdot\mathbf{P}(A\mid Z_i)}=\frac{\frac{30}{100}\cdot\frac{10}{100}}{\frac{1450}{10000}}=0.207$$

٢- في مطبعة للكتب توجد أربع آلات للإنتاج L_1، L_2، L_3 و L_4 لها القدرة الانتّاجية نفسها، ولكن نسبة النسخ المعيبة (غير محقّقة للمواصفات) في إنتاج هذه الآلات هو 0.01 ، 0.03 ، 0.02 و0.05 على الترتيب. نقوم بسحب عشوائي لنسخة كتاب من الإنتاج الكليِّ للمطبعة، فعندئذٍ:

أ – ما هو احتمال أن تكون نسخة الكتاب المسحوبة سليمة؟

ب- إذا وجدنا أنَّ نسخة الكتاب المسحوبة معيبة، فما هو احتمال أن تكون من إنتاج الآلة L_4؟

الحل: للإجابة على هذه الأسئلة سنفترض أنَّ Z_1، Z_2، Z_3 و Z_4 هو الحادث الذي يعبِّر عن كون نسخة الكتاب المسحوبة من إنتاج الآلة L_1، L_2، L_3 و L_4 على الترتيب، فنجد أنَّ هذه الحوادث تشكل تجزئة للحادث الأكيد (الذي يمثّل كل انتاج المطبعة من الكتب)، وبما أنَّ لجميع الآلات القدرة الانتاجية نفسها فإنَّه سيكون لدينا أنَّ:

$$\mathbf{P}(Z_1)=\mathbf{P}(Z_2)=\mathbf{P}(Z_3)=\mathbf{P}(Z_4)=\frac{25}{100}$$

ولذلك فمن أجل الطلب:

أ –سنفترض أنَّ:

- A هو حادث سحب نسخة كتاب معيبة من الإنتاج الكلي للمطبعة.
- B هو حادث سحب نسخة كتاب سليمة من الإنتاج الكلي للمطبعة.

فعندئذٍ يكون لدينا $\mathbf{P}(B)=1-\mathbf{P}(A)$ ، ولكن من صيغة الاحتمال التام لدينا:

$$\mathbf{P}(A)=\sum_{i=1}^{4}\mathbf{P}(Z_i)\cdot\mathbf{P}(A\mid Z_i)$$
$$=\frac{25}{100}\cdot\frac{1}{100}+\frac{25}{100}\cdot\frac{3}{100}+\frac{25}{100}\cdot\frac{2}{100}+\frac{25}{100}\cdot\frac{5}{100}=\frac{275}{10000}=0.275$$

ومن ثمَّ يكون:

$$\mathbf{P}(B) = 1 - \mathbf{P}(A) = 1 - 0.275 = 0.725$$

ب- نجد من صيغة بييز في الاحتمالات أنَّ:

$$\mathbf{P}(Z_4 \mid A) = \frac{\mathbf{P}(Z_4) \cdot \mathbf{P}(A \mid Z_4)}{\sum_{i=1}^{4} \mathbf{P}(Z_i) \cdot \mathbf{P}(A \mid Z_i)} = \frac{\frac{25}{100} \cdot \frac{5}{100}}{\frac{275}{10000}} = 0.4\overline{5}$$

٥ - ٧
استقلال الحوادث

إنَّ مفهوم استقلال الحوادث يُنظر إليه كأحد المفاهيم المهمَّة في الاحتمالات، ولذلك سنقدِّمه من خلال دراسة مبسطة فقط ولن نخوض في تفصيلات تقع خارج مستوى هذا الكتاب.

ليكن لدينا تجربة عشوائية بفضاء حوادث ابتدائية Ω منتهٍ، ولنأخذ A وB حادثين من 2^{Ω}، فإذا كان تَحقُّق وقوع الحادث B لا يؤثر في تَحقُّق وقوع الحادث A ولا بأي شكل من الأشكال، فعندئذٍ يقال إنَّ الحادث A **مستقل** عن الحادث B (**أو الحادث B مستقل عن الحادث A**)، وهذا يعني أنَّه بفرض $0 < \mathbf{P}(B)$ فإنَّ $\mathbf{P}(A \mid B) = \mathbf{P}(A)$، ومن ثمَّ باستخدام **قانون الضرب في الاحتمالات** يكون لدينا الآتي محققاً:

$$\mathbf{P}(A \cap B) = \mathbf{P}(B) \cdot \mathbf{P}(A \mid B) = \mathbf{P}(A) \cdot \mathbf{P}(B)$$

ومنه ينتج أنَّه إذا كان الحادث A مستقل عن الحادث B فإنَّ العلاقة الآتية ستكون مُحقَّقة:

$$\mathbf{P}(A \cap B) = \mathbf{P}(A) \cdot \mathbf{P}(B)$$

وبشكل مماثل يكون B **مستقل** عن A بحسب المفهوم السابق إذا تَحقَّقت العلاقة السابقة أيضاً، ونقول حينئذٍ إنَّ الحادثين A وB مستقلان بعضهما عن البعض الآخر، وهكذا يمكننا أن نصيغ تعريف الاستقلال لحادثين على النحو الآتي.

٥-٧-١- تعريف (استقلال حادثين)

يُقال عن حادثين A وB من 2^{Ω} إنَّهما مستقلان بعضهما عن البعض الآخر إذا تَحقَّقت من أجلهما العلاقة الآتية:

$$\mathbf{P}(A \cap B) = \mathbf{P}(A) \cdot \mathbf{P}(B) \qquad [5\text{-}16]$$

٥-٧-٢- ملاحظات

١- إنَّ الحادث الأكيد Ω مستقل عن أي حادث آخر A من 2^{Ω}، وذلك لأنَّه من أجل حادث A كيفي من 2^{Ω} لدينا العلاقة الآتية محقَّقةً دوماً:

$$\mathbf{P}(A \cap \Omega) = \mathbf{P}(A)$$

$$\mathbf{P}(A) \cdot \mathbf{P}(\Omega) = \mathbf{P}(A) \cdot (1) = \mathbf{P}(A)$$

ومن ثمَّ يكون لدينا:

$$\mathbf{P}(A \cap \Omega) = \mathbf{P}(A) \cdot \mathbf{P}(\Omega)$$

٢- بشكل مماثل لما سبق يمكننا صياغة الاستقلال بين ثلاثة حوادث A وB وC من 2^{Ω} حيث يُقال عن هذه الحوادث إنَّها **مستقلَّة** (**أو مستقلَّة عشوائياً Stochastic Independent**) إذا تَحقَّق من أجلها **جميع** العلاقات الآتية:

$$\mathbf{P}(A\cap B)=\mathbf{P}(A)\cdot\mathbf{P}(B)$$
$$\mathbf{P}(A\cap C)=\mathbf{P}(A)\cdot\mathbf{P}(C)$$
$$\mathbf{P}(B\cap C)=\mathbf{P}(B)\cdot\mathbf{P}(C)$$
$$\mathbf{P}(A\cap B\cap C)=\mathbf{P}(A)\cdot\mathbf{P}(B)\cdot\mathbf{P}(C)$$

٣- إذا كانت العلاقات الثلاث الأولى في الفقرة السابقة مُحقَّقة، فعندئذٍ يُقال عن الحوادث A وB وC إنَّها **مستقلَّة مثنى مثنى** Pairwise Independent، وبناءً على ذلك، فإذا كانت الحوادث التي قيد الدراسة ليست مستقلَّة مثنى مثنى فإنَّها لن تكون مستقلَّة.

٤- إذا كانت الحوادث التي قيد الدراسة مستقلَّة فإنَّها ستكون مستقلَّة مثنى مثنى، ولكنَّ العكس ليس صحيحاً.

◀ ٥-٧-٣- أمثلة

١- لنأخذ تجربة إلقاء قطعة نقود متوازنة لثلاث مرات متتالية، وليكن A وB وC هو حادث الحصول على صورة في الرمية الأولى والثانية والثالثة على الترتيب، فعندئذٍ نجد أنَّ هذه الحوادث مستقلَّة، وذلك لأنَّه لدينا:

$$\Omega=\{HHH,HHT,HTH,THH,TTH,THT,HTT,TTT\}$$
$$A=\{HHH,HHT,HTH,HTT\}$$
$$B=\{HHH,HHT,THH,THT\}$$
$$C=\{HHH,HTH,THH,TTH\}$$

ومن ثمَّ يكون لدينا:

$$\mathbf{P}(A)=\mathbf{P}(B)=\mathbf{P}(C)=\frac{1}{2}$$

$$\mathbf{P}(A\cap B)=\frac{1}{4}=\mathbf{P}(A)\cdot\mathbf{P}(B)$$

$$\mathbf{P}(A\cap C)=\frac{1}{4}=\mathbf{P}(A)\cdot\mathbf{P}(C)$$

$$\mathbf{P}(B\cap C)=\frac{1}{4}=\mathbf{P}(B)\cdot\mathbf{P}(C)$$

$$\mathbf{P}(A\cap B\cap C)=\frac{1}{8}=\mathbf{P}(A)\cdot\mathbf{P}(B)\cdot\mathbf{P}(C)$$

٢- لنفترض أنَّه لدى طالب مقرَّر دراسيّ، ويتقدَّم باختبارين لهذا المقرَّر، تحريري أولاً ومن ثمَّ مقابلة، ويمكن له أن يحصل على أحد تقديرين A أو B بالاحتمال نفسه في أي اختبار من الاختبارات التي سيجريها. عندئذٍ سيكون لفضاء الحوادث الابتدائية الخاص بهذه المسألة العرض الآتي:

$$\Omega = \{AA\ ,\ AB\ ,\ BA\ ,\ BB\}$$

فإذا أخذنا الحوادث الآتية:

$$A_1 = \{AA, AB\} \qquad \& \qquad A_2 = \{AA, BA\} \qquad \& \qquad A_3 = \{AA, BB\}$$

فهل هذه الحوادث مستقلّة بعضها عن البعض الآخر؟

الحل: للإجابة على هذا السؤال لدينا من معطيات المسألة ما يلي:

$$\mathbf{P}(\{AA\}) = \mathbf{P}(\{AB\}) = \mathbf{P}(\{BA\}) = \mathbf{P}(\{BB\}) = \frac{1}{4}$$

$$\mathbf{P}(A_1) = \mathbf{P}(A_2) = \mathbf{P}(A_3) = \frac{1}{2}$$

$$\mathbf{P}(A_1 \cap A_2) = \mathbf{P}(A_1) \cdot \mathbf{P}(A_2) = \frac{1}{4}$$

$$\mathbf{P}(A_1 \cap A_3) = \mathbf{P}(A_1) \cdot \mathbf{P}(A_3) = \frac{1}{4}$$

$$\mathbf{P}(A_2 \cap A_3) = \mathbf{P}(A_2) \cdot \mathbf{P}(A_3) = \frac{1}{4}$$

وهذا يعني أنَّ الحوادث A_1 و A_2 و A_3 مستقلَّة مثنى مثنى، ولكنَّ هذه الحوادث ليست مستقلَّة عشوائياً وذلك لأنَّ:

$$\mathbf{P}(A_1 \cap A_2 \cap A_3) = \mathbf{P}(\{AA\}) = \frac{1}{4} \neq \mathbf{P}(A_1) \cdot \mathbf{P}(A_2) \cdot \mathbf{P}(A_3) = \frac{1}{8}$$

٣- لنأخذ تجربة إلقاء حجر نرد متوازن لمرتين متتاليتين فقط، ولنأخذ الحوادث الآتية:

$$A = \left\{ \begin{matrix} (1,1)\ ,(1,2)\ ,(1,5)\ (2,1),(2,2),(2,5), \\ (3,1),(3,2),(3,5),(4,1),(4,2),(4,5), \\ (5,1),(5,2),(5,5),(6,1),(6,2),(6,5) \end{matrix} \right\}$$

$$B = \left\{ \begin{matrix} (1,4)\ ,(1,5)\ ,(1,6),(2,4),(2,5),(2,6), \\ (3,4),(3,5),(3,6),(4,4),(4,5),(4,6), \\ (5,4),(5,5),(5,6),(6,4),(6,5),(6,6) \end{matrix} \right\}$$

$$C = \{(3,6),(4,5),(5,4),(6,3)\}$$

فنجد أنَّ هذه الحوادث ليست مستقلَّة مثنى مثنى وذلك لأنَّ:

$$\mathbf{P}(A \cap B) = \frac{1}{6}$$

$$\mathbf{P}(A)\cdot\mathbf{P}(B) = \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4}$$

علماً أنّه لدينا:

$$\mathbf{P}(A \cap B \cap C) = \frac{1}{36} = \mathbf{P}(A)\cdot\mathbf{P}(B)\cdot\mathbf{P}(C)$$

٥-٧-٤- بعض خصائص الحوادث المستقلّة

فيما يلي نقدّم بعض الخصائص للحوادث المستقلّة دون الخوض في إثباتاتها.

ليكن A وB حادثين مستقلّين بعضهما عن البعض الآخر من فضاء حوادث ابتدائية Ω، فعندئذٍ:

١- يكون الحادثان A و $\bar{B}$ مستقلين بعضهما عن البعض الآخر أيضاً.

٢- يكون الحادثان $\bar{A}$ و B مستقلين بعضهما عن البعض الآخر أيضاً.

٣- يكون الحادثان $\bar{A}$ و $\bar{B}$ مستقلين بعضهما عن البعض الآخر أيضاً.

٥-٧-٥- مثال

لدينا ثلاث محطّات لتوليد الطاقة الكهربائية E_1، E_2 و E_3 وتعمل كل منها بشكل مستقلٍّ عن المحطّتين الأخريين، فإذا كان احتمال تعطّل هذه المحطّات خلال السّنة القادمة هو 0.07 ، 0.05 و0.03 على الترتيب، فعندئذٍ:

أ- ما هو احتمال عمل المحطّات الثلاث خلال السّنة القادمة؟

ب- ما هو احتمال عمل المحطّة E_2 وتعطّل E_1 و E_3 خلال السّنة القادمة؟

ج- ما هو احتمال تعطّل محطّةٍ واحدةٍ على الأكثر خلال السّنة القادمة؟

الحل: للإجابة على هذه الأسئلة سنفترض أنَّ A_1، A_2 و A_3 هو حادث تعطّل المحطّة E_1، E_2 و E_3 خلال السّنة القادمة على الترتيب. عندئذٍ بحسب الفرض ستكون الحوادث A_1 و A_2 و A_3 مستقلّة عشوائياً، ومنه يكون لدينا من أجل الطلب:

أ- بفرض أنَّ A هو الحادث الذي يُعبِّر عن عمل المحطّات الثلاث خلال السّنة القادمة، فإنَّه سيكون لدينا $A = \bar{A}_1 \cap \bar{A}_2 \cap \bar{A}_3$، ومن ثمَّ بسبب استقلال الحوادث A_1 و A_2 و A_3 يمكننا أن نكتب الآتي:

$$\mathbf{P}(A) = \mathbf{P}(\bar{A}_1 \cap \bar{A}_2 \cap \bar{A}_3) = \mathbf{P}(\bar{A}_1)\cdot\mathbf{P}(\bar{A}_2)\cdot\mathbf{P}(\bar{A}_3)$$
$$= 0.93 \times 0.95 \times 0.97 = 0.857$$

ب- بفرض أنَّ B هو الحادث الذي يُعبِّر عن عمل المحطّة E_2 وتعطّل E_1 و E_3 خلال السّنة القادمة، فإنَّه سيكون لدينا $B = A_1 \cap \overline{A_2} \cap A_3$، ومن ثمَّ بسبب استقلال الحوادث A_1 و A_2 و A_3 يمكننا أن نكتب:

$$\mathbf{P}(B)=\mathbf{P}(A_1\cap\bar{A}_2\cap A_3)=\mathbf{P}(A_1)\cdot\mathbf{P}(\bar{A}_2)\cdot\mathbf{P}(A_3)$$
$$=0.07\times0.95\times0.03=0.002$$

ج- بفرض أنَّ C هو الحادث الذي يُعبِّر عن تعطّل محطّةٍ واحدةٍ على الأكثر خلال السّنة القادمة، فعندئذٍ يكون لدينا:

$$C=\left(\bar{A}_1\cap\bar{A}_2\cap\bar{A}_3\right)\cup\left(\bar{A}_1\cap\bar{A}_2\cap A_3\right)\cup\left(\bar{A}_1\cap A_2\cap\bar{A}_3\right)\cup\left(A_1\cap\bar{A}_2\cap\bar{A}_3\right)$$

هو الحادث الذي يُعبِّر عن تعطّل محطّة واحدةٍ على الأكثر خلال السّنة القادمة، فعندئذٍ بملاحظة أنَّ الحوادث:

$$\bar{A}_1\cap\bar{A}_2\cap\bar{A}_3 \quad \& \quad \bar{A}_1\cap\bar{A}_2\cap A_3 \quad \& \quad \bar{A}_1\cap A_2\cap\bar{A}_3 \quad \& \quad A_1\cap\bar{A}_2\cap\bar{A}_3$$

متنافية مثنى مثنى، وبسبب استقلال الحوادث A_1 و A_2 و A_3، ومن ثمَّ استخدام خصائص الحوادث المستقلّة، فإنَّه يمكننا أن نكتب الآتي:

$$\mathbf{P}(C)=\mathbf{P}(\bar{A}_1)\cdot\mathbf{P}(\bar{A}_2)\cdot\mathbf{P}(\bar{A}_3)+\mathbf{P}(\bar{A}_1)\cdot\mathbf{P}(\bar{A}_2)\cdot\mathbf{P}(A_3)$$
$$+\mathbf{P}(\bar{A}_1)\cdot\mathbf{P}(A_2)\cdot\mathbf{P}(\bar{A}_3)+\mathbf{P}(A_1)\cdot\mathbf{P}(\bar{A}_2)\cdot\mathbf{P}(\bar{A}_3)$$
$$=0.97\times0.95\times0.93+0.97\times0.95\times0.03$$
$$+0.97\times0.05\times0.93\times+0.03\times0.95\times0.97=0.99311$$

▶



١- لدينا باب خزنة حديدية مقفل ويُفتح بنظام رقمي عشري مكوَّن من خمس خانات. عندئذٍ:

أ- إذا كان من الممكن تكرار الرقم، فما هو أكبر عدد من المحاولات اللازمة لفتح هذا الباب؟

ب- إذا كان من غير الممكن تكرار الرقم، فما هو أكبر عدد من المحاولات اللازمة لفتح هذا الباب؟

٢- بكم طريقة يمكن لستة طلاب الجلوس:

أ- على مقعد خشبي مفتوح يتسع لستة أشخاص على الأقل في فصل دراسي؟

ب- على طاولة مستديرة حولها ستة كراسي في فصل دراسي؟

٣- إذا كانت المدينة **A** ترتبط بالمدينة **B** بخمس طرق مختلفة، وكانت المدينة **B** ترتبط بالمدينة **C** بثلاث طرق مختلفة. بكم طريقة يمكن لشخص السفر من المدينة **A** إلى المدينة **C** مروراً بالمدينة **B**؟

٤- تريد مؤسسة تشكيل لجنة مكوَّنة من خمسة خبراء لوضع خطتها الاستراتيجية للأعوام الخمس القادمة. بكم طريقة يمكن تشكيل هذه اللجنة إذا علمت أنَّ:

أ- المؤسسة تحوي 15 خبيراً لهم جميعاً الكفاءة نفسها (أي يمكن لكل منهم أن يقوم بأية مهمّة توكل إليه)؟

ب- المؤسسة تحوي 15 خبيراً وسوف توكل لكلِّ واحدٍ منهم مهمّة خاصّة به؟

٥- يوجد في إدارة منظّمّة اجتماعية تسعة رجال وثمان نساء، وطلب منهم تشكيل وفد لمناظرة تلفزيونية حول هذه المنظّمّة وعلى أن يكون الوفد مكوَّن من رجلين وثلاث نساءٍ أو ثلاثة رجالٍ وامرأتين. بكم طريقة يمكن اختيار هذا الوفد؟

٦- يوجد على قائمة الطعام في مطعم خمسة أصناف من المقبلات واثنتي عشرة صنفاً من الأطعمة وستة أصناف من المشروبات الغازية، بكم طريقة يمكن لزبون أن يختار صنفين من المقبلات وثلاثة أنواعٍ من الأطعمة ومشروبٍ غازيٍّ واحد؟

٧- يوجد في صندوقٍ كرتين بيضاويين، كرتين سوداويين وأربع كرات زرقاء. نقوم بسحب ثلاث كراتٍ على التوالي مع الإرجاع (أو الإعادة). عندئذٍ:

أ- إذا كان الحادث A هو الحصول على ثلاثة كرات زرقاء فما هو الحادث $\bar{A}$؟

ب- إذا كان الحادث B هو حادث الحصول على كرة سوداء في السحب الأول وبيضاء في السحب الثاني، وبفرض أنَّ C هو حادث الحصول على كرة زرقاء في السحب الثالث، فما هو الحادث $B \cup C$؟

ج- إذا كان الحادث D هو حادث الحصول على كرة سوداء في السحب الأول، وبفرض أنَّ E هو حادث الحصول على كرة زرقاء في السحب الثالث، فما هو الحادث $D \cap E$؟

٨- لدى أسرة ثلاثة أطفال، فإذا كان لكلٍ منهم النصيب نفسه في الوجود (الولادة)، فما هو احتمال أن يكون:

أ- لدى الأسرة ولدين وبنت؟

ب- لدى الأسرة ولد واحد على الأقل؟

٩- يوجد في قسم تسعة أعضاء هيئة تدريس منهم ثلاث إناث، وطُلِبَ من القسم تشكيل لجنة علميّة من أعضائه مكوَّنة من أربعة أشخاص. فما هو احتمال:

أ- أن تكون جميع الإناث في هذه اللجنة؟

ب- أن يكون في هذه اللجنة أنثى واحدة على الأقل؟

ج- أن يكون في هذه اللجنة رجل واحد على الأقل؟

١٠- خصَّص مدرب 30 سؤالاً لاختبار الطلاب، وبحيث يقدَّم لكل طالب خمسة أسئلة تسحب عشوائياً من هذه الأسئلة، فإذا علمت أنَّه يمكن لطالب X أن يجيب على 16 سؤال من الــ 30 المخصَّصة للاختبار، فما هو احتمال أن يجيب الطالب على الأسئلة الخمسة في الامتحان؟

١١- لنأخذ تجـربة قذف قطعة نقود متـوازنة مع إلقاء حجر نرد متوازن لمرّة واحدة. عندئذٍ عيّن فضاء الحوادث الابتدائية لهذه التجربة العشوائية، ومن ثمَّ اجب عما يلي:

أ- احسب احتمال الحصول على صورة وعدد فردي.

ب – احسب احتمال الحصول على شعار وعدد أكبر من 5.

ج- هل حادث الحصول على صورة وعدد فردي مستقل عن حادث الحصول على شعار وعدد زوجيٍّ؟

١٢- لنأخذ تجربة إلقاء حجر نرد متوازن لمرتين متتاليتين، فعندئذٍ:

أ- احسب احتمال أن يكون مجموع الرقمين الظاهرين للأعلى أصغر من 2.

ب- احسب احتمال أن يكون مجموع الرقمين الظاهرين للأعلى أصغر من 13.

جـ – احسب احتمال أن يكون الرقمين الظاهرين للأعلى غير متساويين.

د- ما هو احتمال الحصول على رقمين متتاليين؟

هـ- هل حادث الحصول على الرقم 1 في الرمية الأولى مستقل عن حادث الحصول على الرقم نفسه في الرمية الثاني؟

و- لو أخذنا تجربة إلقاء حجري نرد متوازنين ومتمايزين لمرّة واحدة فقط، فهل تتغيرٍّ نتائج الطلبات السابقة في هذه المسألة، ولماذا؟

١٣- لنأخذ تجربة إلقاء حجري نردٍ متوازنين ومتماثلين تماماً (غير متمايزين) لمرّة واحدة، فعندئذٍ اجب عن الأسئلة الآتية:

أ- احسب احتمال أن يكون مجموع الرقمين الظاهرين للأعلى أكبر من 6.

ب- احسب احتمال أن يكون الرقمين الظاهرين للأعلى متساويين.

ج- ما هو احتمال الحصول على رقمين غير متتاليين؟

د- هل حادث الحصول على الرقم 3 في الرمية الأولى مستقلٌّ عن حادث الحصول على الرقم 2 في الرمية الثانية؟

١٤- في صندوق يوجد 6 كرات متماثلة تماماً، منها ثلاث كرات حمراء وكرتان زرقاوين والكرة السادسة خضراء. قمنا بسحب عشوائي لثلاث كرات بآن واحد من هذا الصندوق، والمطلوب:

أ- ما هو احتمال أن تكون الكرات الثلاث من ألوان مختلفة؟

ب- ما هو احتمال أن تكون الكرات الثلاث من لون واحد؟

ج- ما هو احتمال أن يكون لكرتين اللون نفسه والثالثة من لون آخر؟

١٥- يوجد في فصلٍ دراسيٍّ لطلاب السّنة الأولى المشتركة 21 طالباً منهم خمسة طلاب متميّزين، وأردنا تشكيل فريق عمل مُكوَّن من ثلاثة طلاب لإدارة شؤون الفصل، فإذا أخذنا عشوائياً ثلاثة طلاب دفعة واحدة من هذا الفصل، فعندئذٍ:

أ- ما هو احتمال أن يكون في هذا الفريق طالب متميّز واحد فقط؟

ب- ما هو احتمال أن يكون جميع عناصر الفريق من الطلاب المتميّزين؟

١٦- لنفرض أنّـه لدينا صندوقان I وII يحويان كرات متماثلة بألوان مختلفة، وبحيث يحوي الصندوق الأول I على كرتين بيضاويين وكرة سوداء بينما يحوي الصندوق الثاني II كرة بيضاء وكرة سوداء وكرة زرقاء. الآن نقوم بخلط الكرات في الصندوق الأول جيداً ومن ثمَّ نسحب (سحب عشوائي) كرة من هذا الصندوق ونضعها في الصندوق الثاني دون رؤيتها، وبعد ذلك نقوم بخلط الكرات في الصندوق الثاني جيداً ومن ثمَّ نسحب كرة من هذا الصندوق. عندئذٍ لنحسب احتمال أن تكون الكرة المسحوبة أخيراً ذات لون أسود.

١٧- لدينا ثلاثة صناديق بحيث يحوي الصندوق الأول أربع كرات حمراء وثلاث كرات سوداء بينما يحوي الثاني على ثلاث كرات حمراء وخمس كرات بيضاء، وأخيراً يحوي الصندوق الثالث على ثلاث كرات سوداء وثلاث كرات بيضاء. الآن بفرض أنَّ لكل صندوق النصيب نفسه في السحب وأنَّ جميع الكرات **متماثلة** تماماً (**ومن ثمَّ لها النصيب نفسه في الاختيار**)، وأنّنا قمنا بسحب عشوائي لصندوق من هذه الصناديق، ومن ثمَّ سحب كرة من هذا الصندوق، فعندئذٍ:

أ – ما هو احتمال أن تكون الكرة المسحوبة بيضاء؟

ب- إذا علمنا أنَّ الكرة المسحوبة كانت بيضاء، فما هو احتمال أن تكون هذه الكرة قد سُحِبت من الصندوق الثاني؟

١٨- في مصنع لإنتاج المصابيح الكهربائية يوجد ثلاثة خطوط للإنتاج L_1 ، L_2 و L_3 بحيث تنتج هذه الخطوط 50 % ، 30 % و 20 % على الترتيب من نسبة الإنتاج الكلي للمصنع، ونسبة المعطَّل في إنتاج هذه الخطوط على الترتيب هو 0.03 ، 0.02 و0.01 . الآن نقوم بسحبٍ عشوائيٍّ لمصباحٍ من الإنتاج الكليِّ للمصنع، فما هو احتمال أن يكون:

أ – المصباح المسحوب سليماً؟

ب- المصباح المسحوب من إنتاج الخط الأول إذا علمنا أنَّه وجد مُعطَّلاً؟

١٩- إذا كان A وB حادثين من فضاء حوادث ابتدائية Ω ، علماً أنَّ:

$$P(A)=0.4 \quad , \quad P(B)=0.5 \quad \text{and} \quad P(A \cap B)=0.4$$

عندئذٍ أيّ من القيم الآتي هي $P(A \cup B)$ ؟

A) 0.3 B) 0.2 C) 0.5 D) 0.1

٢٠- إذا كان $\Omega=\{E_1, E_2, E_3\}$ فضاء حوادث ابتدائية، وكان لدينا $P(E_1)=P(E_3)=0.30$ ، فعندئذٍ أيٌّ من القيم الآتية تساوي $P(E_2)$ ؟

A) 0.04 B) 0.40 C) 0.41 D) 0.30

٢١- إذا كان A و B حادثين من فضاء حوادث ابتدائية Ω بحيث أنَّ $P(A)=0.6$ و $P(B \cap A)=0.3$. عندئذٍ أي من القيم الآتية تساوي $P(A \setminus B)$ ؟

A) 0.30 B) 0.03 C) 0.50 D) 0.05

٢٢- إذا كان A و B حادثين مستقلَّين من فضاء حوادث ابتدائية Ω بحيث أنَّ $P(A)=0.4$ و $P(B)=0.1$. عندئذٍ أي من القيم الآتية تساوي $P(A \cap B)$ ؟

A) 0.04 B) 0.40 C) 0.10 D) 0.004

٢٣- إذا كان A و B حادثين من فضاء حوادث ابتدائية Ω بحيث أنَّ $P(A)=P(B)=0.495$ و $P(A \cap B)=0.09$. عندئذٍ أي من القيم الآتية تساوي $P(A \cup B)$ ؟

A) 0.50 B) 0.51 C) 0.90 D) 0.001

٢٤- صندوق يحتوي على 15 كرة متماثلة تماماً منها 5 حمراء، 4 زرقاء، و6 خضراء. إذا سحبت كرة بشكل عشوائي من الصندوق، فما هو احتمال أن يكون لونها أزرق أو أحمر؟

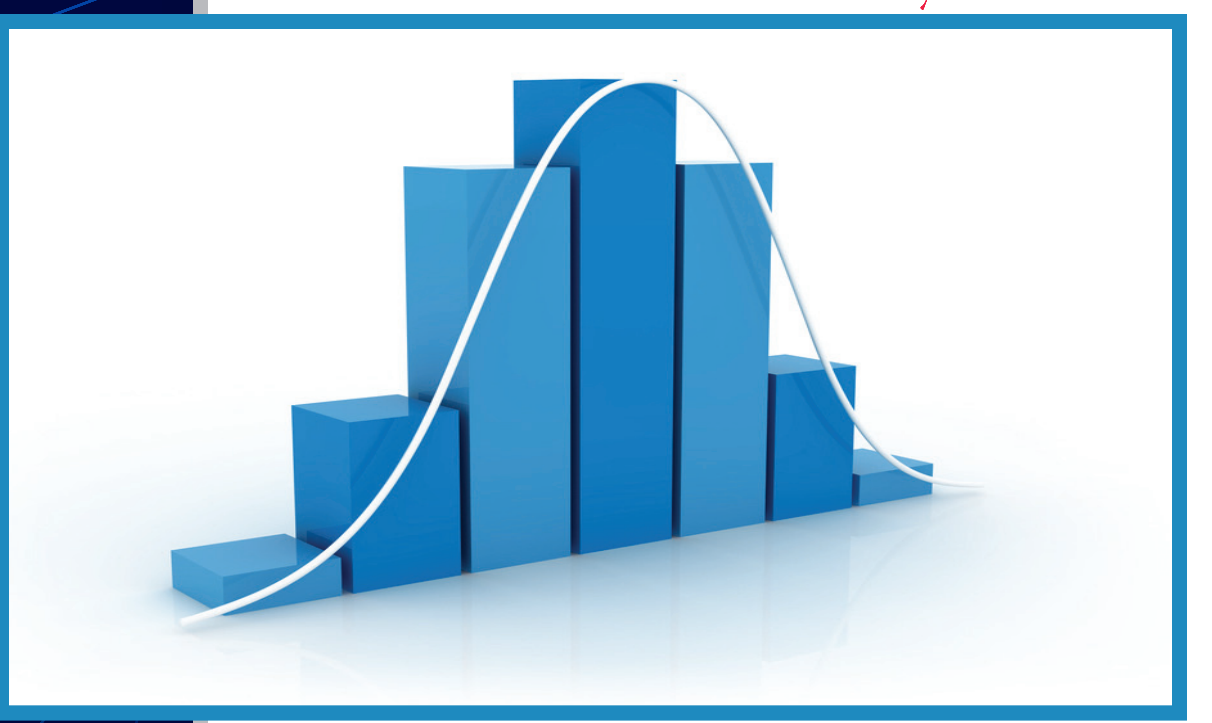
A) 4/15 B) 6/15 C) 9/15 D) 5/15

٢٥- إذا استخدمت جميع حروف OKLAH في تكوين كلمة، فكم كلمة يمكن تكوينها من هذه الأحرف؟

A) 100 B) 120 C) 80 D) 166

الفصل السادس

المتغيّرات العشوائية وتوزيعاتها الاحتمالية
Random Variable and their Probability Distributions



المقدمة:

لقد قدَّمنا في الفصل السابق مفهوم الدالّة الاحتماليّة وبعض خصائصها، ومن ثمَّ تناولنا كيفية حساب احتمالات بعض الحوادث المتعلّقة بتجربة عشوائيّة ذات عدد منتهٍ من النتائج، وبعد ذلك تطرقنا (وبشكل مبسَّط) إلى الاحتمال الشرطي وبعض خصائصه، وأخيراً خُتِمَ الفصل بعرض مفهوم الاستقلال لعددٍ منتهٍ من الحوادث (حتى ثلاثة حوادث). لكن يصادفنا في كثير من مسائل الاحتمالات طروحات تجعل من اهتمامنا بالحادث الابتدائيّ منفرداً ليس مرغوباً ويكون تركيزنا منصباً على الخصائص التي يحملها الحادث الابتدائيّ، بمعنى أنَّنا لا نّهتم أيّ حادثٍ ابتدائيّ سنأخذ وإنّما المهمّ إن كان الحادث الابتدائيّ يتمتّع بصفات معيّنة أم لا. إنَّ هذه المسائل تترافق عادةً مع مفهوم المتغيِّر العشوائيّ والذي سيكون محور دراستنا لهذا الفصل.

- ٦ - ١ - المتغيِّرات العشوائية
- ٦ - ٢ - دالَّة التوزيع لمتغيِّر عشوائي
- ٦ - ٣ - المتغيِّرات العشوائية المتقطِّعة
- ٦ - ٤ - المتغيّرات العشوائية المستمرّة

٦ - ١
المتغيِّرات العشوائية

لتوضيح مفهوم المتغيِّر العشوائيّ سنقدِّم المثال الآتي.

◀ ٦-١-١- مثال

لنأخذ تجربة قذف قطعة نقود لمرّتين متتاليتين، فعندئذٍ سيكون لفضاء الحوادث الابتدائيّة لهذه التجربة العشوائيّة العرض الآتي:

$$\Omega = \{HH, HT, TH, TT\}$$

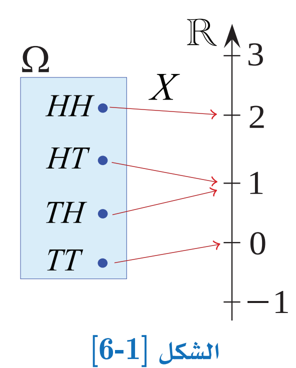
فلو افترضنا أنَّ A هو حادث حصولنا على صورةٍ واحدةٍ على الأقلّ خلال الرميتين، فعندئذٍ يكون:

$$A = \{HH, HT, TH\}$$

حيث نلاحظ أنَّه لا أهمية لتوضّعات الصور خلال الرميتين، وهذا يعني أنَّنا لم نعد نّهتم في أيّة رميّة من هاتين الرميتين سنحصل على الصورة وإنّما الذي يهمنا حقاً هو إن كنَّا سنحصل في نهاية الرميتين على صورةٍ واحدةٍ على الأقل، ولذلك دراسة كلّ حادث ابتدائيّ ω من Ω لم تعدْ مهمّة بالنسبة لنا في هذه التجربة.

الآن لو أمعنّا النظر في مكوّنات الحادث A، فإنَّه قد يتبادر إلى ذهننا فكرة استخدام الأعداد للتعبير عن الحادث A وذلك من خلال عمليّة إرفاق كلّ حادث ابتدائيّ ω من Ω بعددٍ يعبرّ عن عدد الصور في هذا الحادث الابتدائيّ (انظر الشكل [6-1]).

إنَّ عملية الإرفاق هذه ما هي إلاَّ تطبيق X مُعرَّف على Ω (مجاله Ω) ويأخذ قيمه في مجموعة الأعداد الحقيقية $\mathbb{R}$ (مجاله المقابل $\mathbb{R}$) كما يوضِّحه الشكل [6-1]، وأكثر من ذلك فإنَّ هـذا التطبيق وحيد القيمة ومعيَّن بشكل وحيد وفقاً للطريقة التي قُدِّم بها، وهذا يعني أنَّ التطبيق X هو دالّة حقيقية مُعرَّفة على Ω.



الشكل [6-1]

الآن لو رمزنا لمجموعة قيم التطبيق X بـ Ω^*، فعندئذٍ سيكون لدينا $\Omega^* = \{0,1,2\}$، ونلاحظ هنا أنَّ كلّ عنصر من المجموعة Ω^* هو من جديد حادث ابتدائيّ لأن ظهور القيم 0 و1 و2 في Ω^* سيكون عشوائياً، وعشوائيّتها تنتج من عشوائيّة الحوادث الابتدائيّة في Ω.

في الحقيقة يوجد **شرط هام جداً** يجب تَحقيقه من قِبَل التطبيق X، وبدونه لا يمكن القبول بعشوائيّة هذا التطبيق أبداً. إنَّ هذا الشرط ينصّ على أن تكون الصورة العكسيّة لأيّ حادثٍ مكوَّن من عناصر Ω^* هو حادثٌ في Ω أيضاً، وهذا يعني أنَّه من أجل أي حادث B من 2^{Ω^*} فإنَّ $X^{-1}(B)$ يجب أن يكون حادثاً من 2^{Ω}.

إذاً لو أردنا التّحقّق من عشوائيّة التطبيق X في مثالنا المقدَّم أعلاه علينا إثبات أنَّه من أجل أي حادثٍ B من 2^{Ω^*} فإنَّ $X^{-1}(B)$ هو حادثٌ من 2^{Ω}، وللبحث في ذلك لدينا من الفصل السّابق:

$$2^{\Omega}=\left\{\begin{array}{l}\varnothing, \{HH\}, \{HT\},\{TH\}, \{TT\},\{HH,HT\}, \{HH,TH\},\\ \{HH,TT\},\{HT,TH\}, \{HT,TT\}, \{TH,TT\}, \{HH,HT,TH\},\\ \{HH,HT,TT\}, \{HH,TH,TT\},\{HT,TH,TT\},\\ \{HH,HT,TH,TT\}=\Omega\end{array}\right\}$$

ومن أجل Ω^* نجد أنَّ:

$$2^{\Omega^*}=\left\{\varnothing, \{0\}, \{1\},\{2\}, \{0,1\}, \{0,2\},\{1,2\},\{0,1,2\}=\Omega\right\}$$

ومن ثمَّ يكون:

من أجل $B=\varnothing$ لدينا: $X^{-1}\left(\varnothing\right)=\varnothing\in 2^{\Omega}$

من أجل $B=\{0\}$ لدينا: $X^{-1}\left(\{0\}\right)=\{TT\}\in 2^{\Omega}$

من أجل $B=\{1\}$ لدينا: $X^{-1}\left(\{1\}\right)=\{HT,TH\}\in 2^{\Omega}$

من أجل $B=\{2\}$ لدينا: $X^{-1}\left(\{2\}\right)=\{HH\}\in 2^{\Omega}$

من أجل $B=\{0,1\}$ لدينا: $X^{-1}\left(\{0,1\}\right)=\{TT,HT,TH\}\in 2^{\Omega}$

من أجل $B=\{0,2\}$ لدينا: $X^{-1}\left(\{0,2\}\right)=\{TT,HH\}\in 2^{\Omega}$

من أجل $B=\{1,2\}$ لدينا: $X^{-1}\left(\{1,2\}\right)=\{HT,TH,TT\}\in 2^{\Omega}$

من أجل $B=\{0,1,2\}=\Omega^*$ لدينا: $X^{-1}\left(\Omega^*\right)=\{HH,HT,TH,TT\}=\Omega\in 2^{\Omega}$

وهذا يعني أنَّ الشرط المذكور أعلاه مُحقَّق من أجل التطبيق X، ومن ثمَّ X هو متغيِّر عشوائيّ حقّاً.

▶

تجدر الإشارة هنا إلى أنَّ احتمال أن يأخذ التطبيق X قيمة من قيمه ليست بالضرورة أن تكون متساوية من أجل القيم المختلفة لـ X حتى لو كان لجميع الحوادث الابتدائيّة الاحتمال نفسه، فعلى سبيل المثال لو كانت قطعة النقود **متوازنة** فعندئذٍ يُلاحظ أنَّ احتمال أن يأخذ التطبيق X إحدى قيمه مرتبطة بعدد الحوادث الابتدائيّة النّاتجة عنها في Ω، ومن ثمَّ يكون لدينا ما يلي (**حيث لجميع عناصر Ω النصيب نفسه في الظهور**):

$$\mathbf{P}\left(\{\omega \in \Omega \,;\, X(\omega) = 0\}\right) = \frac{1}{4} = 0.25$$

$$\mathbf{P}\left(\{\omega \in \Omega \,;\, X(\omega) = 1\}\right) = \frac{2}{4} = 0.5$$

$$\mathbf{P}\left(\{\omega \in \Omega \,;\, X(\omega) = 2\}\right) = \frac{1}{4} = 0.25$$

إنَّ التطبيقات (أو الدوال) التي تُحقِّق الخاصية المذكورة سابقاً تُدعى متغيِّرات عشوائيّة، وأصل التسمية تعود إلى كون القيم التي يأخذها التطبيق X هي قيم عشوائيّة، وأمّا التعريف الرياضياتي لهذا المفهوم فتقدِّمه لنا الفقرة الآتية.

٦-١-٢- تعريف المتغيِّر العشوائيّ

لتكن Ω مجموعة نتائج تجربة عشوائيّة (منتهية النتائج)، وليكن X تطبيقاً حقيقياً معرَّفاً على Ω بمجموعة قيم Ω^*، فإذا كان من أجل أي حادث B من 2^{Ω^*} لدينا $X^{-1}(B)$ هو حادث من 2^{Ω}، فعندئذٍ يُقال عن التطبيق X إنَّه متغيِّر عشوائيّ على Ω.

٦-١-٣- ملاحظات

١- في الجوانب التطبيقيّة يُعدّ تحقيق الشرط المذكور في متن التعريف السابق ليس سهلاً في كثير من الحالات، ولذلك قُدّمت اختبارات لتحديد ما إذا كان تطبيق X على Ω هو متغيِّر عشوائيّ أم لا، ومن أجل الحالة التي تكون فيها Ω منتهية يمكن اعتماد الاختبار الآتي:

يكون تطبيق X متغيِّراً عشوائياً على Ω إذا كانت المجموعة $\{\omega \in \Omega \,;\, X(\omega) < x\}$ هي حادث من 2^{Ω} وذلك عندما تمسح x كلّ القيم الممكنة لها في $\mathbb{R}$، فعلى سبيل المثال لو عدنا إلى المثال السابق فإنَّنا نجد ما يلي:

$$\{\omega \in \Omega \,;\, X(\omega) < x\} = \begin{cases} \varnothing & \text{for } x \le 0 \\ \{TT\} & \text{for } 0 < x \le 1 \\ \{TT, HT, TH\} & \text{for } 1 < x \le 2 \\ \{TT, HT, TH, TT\} = \Omega & \text{for } x > 2 \end{cases}$$

حيث نعلم أنَّ $\varnothing$ و Ω هي عناصر من 2^{Ω} دوماً، ولدينا $\{TT\}$ و $\{TT, HT, TH\}$ هي عناصر من 2^{Ω} أيضاً، وهذا يعني أنَّ التطبيق X المُعطى في المثال السابق هو متغيِّر عشوائيّ على Ω.

٢- في إطار دراستنا في هذا الكتاب لن نتطرق إلى تحقيق X للشرط المذكور في متن التعريف السابق وذلك لأنَّنا سنتعامل مع تطبيقات تكون مُحقِّقة لهذا الشرط دائماً.

٦-١-٤- مثال

١- لنأخذ تجربة إلقاء حجر نرد متوازن لمرتين متتاليتين، وليكن X **متغيرًا عشوائياً** يُرفق كلّ حادث ابتدائيّ (i,j) بعدد يساوي مجموع المركّبتين i و j ، ولنقم بالإجابة على الأسئلة الآتية:

أ- ما هو احتمال أن يأخذ المتغيِّر العشوائيّ X القيمة 1 ؟

ب- ما هو احتمال أن يأخذ المتغيِّر العشوائيّ X قيمة أصغر من 4 ؟

ج- ما هو احتمال أن يأخذ المتغيِّر العشوائيّ X قيمة أكبر أو يساوي 10 ؟

الحل: نلاحظ هنا أنَّ مجموعة تعريف المتغيِّر العشوائيّ X هي:

$$\Omega=\begin{Bmatrix}(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),\\(2,1),(2,2),(2,3),(2,4),(2,5),(2,6),\\(3,1),(3,2),(3,3),(3,4),(3,5),(3,6),\\(4,1),(4,2),(4,3),(4,4),(4,5),(4,6),\\(5,1),(5,2),(5,3),(5,4),(5,5),(5,6),\\(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\end{Bmatrix}$$

وأمَّا مجموعة قيمه فهي $\Omega^*=\{2,3,4,5,6,7,8,9,10,11,12\}$، ومن ثمَّ يكون لدينا من أجل الطلب:

أ- الاحتمال المطلوب هو:

$$\mathbf{P}\Big(\{(i,j)\in\Omega \mid X((i,j))=1\}\Big)=\mathbf{P}\Big(\{\}\Big)=\mathbf{P}\Big(\varnothing\Big)=0$$

ب- الاحتمال المطلوب هو:

$$\mathbf{P}\Big(\{(i,j)\in\Omega \mid X((i,j))<4\}\Big)=\mathbf{P}\Big(\{(1,1),(1,2),(2,1)\}\Big)=\frac{3}{36}=0.08\overline{3}$$

ج- الاحتمال المطلوب هو:

$$\mathbf{P}\Big(\{(i,j)\in\Omega \mid X((i,j))\ge 10\}\Big)=$$

$$=\mathbf{P}\Big(\{(4,6),(5,5),(6,4),(6,5),(5,6),(6,6)\}\Big)=\frac{6}{36}=0.1\overline{6}$$

٦-١-٥- ملاحظات

١- على سبيل الاختصار والتبسيط يُكتب في كثير من الحالات:

$\mathbf{P}(X=x)$ عوضاً عن $\mathbf{P}\Big(\{\omega\in\Omega\ ;\ X(\omega)=x\}\Big)$

$\mathbf{P}(X<x)$ عوضاً عن $\mathbf{P}\Big(\{\omega\in\Omega\ ;\ X(\omega)<x\}\Big)$

وبالمثل عند استخدام بقيّة المتباينات الممكنة $\le$، $>$ و$\ge$، وعندما نكتب الصيغة المفصَّلة فإنَّه سيكون من باب التذكير بها أو لضرورة تُمليها طبيعة عرض العلاقة التي قيد الدراسة.

٢- لكي نتجنب سوء الفهم الذي يمكن أن ينشأ عن مصطلح المتغيِّر العشوائيّ كدالَّة نودّ أن نشير هنا إلى أنَّ المتغيِّر العشوائيّ X كدالَّة إنَّما يعني أنَّه لدى متغيِّرات مستقلَّة ω من Ω مُعطاة تكون القيم $X(\omega)$ وحيدة التعيين، وأنَّ قبول هذه الدالّة كمتغيِّر عشوائيّ يتوقف على اختيار المتغيِّرات المستقلَّة ω من Ω.

٦-١-٦- أمثلة

١- ليكن Ω فضاء الحوادث الابتدائيّة لرمي مكعَّب (كحجر النرد) دوِّن على كلّ وجه من وجوهه الستة أحد الأرقام 123، 132، 213، 231، 312 و321، أي أنَّ:

$$\Omega := \{\underset{\omega_1}{123}, \underset{\omega_2}{132}, \underset{\omega_3}{213}, \underset{\omega_4}{231}, \underset{\omega_5}{312}, \underset{\omega_6}{321}\}$$

والآن لنأخذ X تطبيقاً حقيقياً معرَّفاً على Ω بحيث يقرن كلّ حادثٍ ابتدائيّ ω_i بعدد يساوي مجموع الأعداد المكوِّنة لهذا الحادث الابتدائيّ، أي أنَّ:

$$X : \Omega \longrightarrow \mathbb{R} \quad ; \omega_i \mapsto X(\omega_i) = 6 \qquad ; i = 1,2,\cdots,6$$

عندئذٍ نجد أنَّ هذا التطبيق هو متغيِّر عشوائيّ وذلك لأنَّ:

$$\{\omega \in \Omega \mid X(\omega) < x\} = \begin{cases} \varnothing & for \quad x \le 6 \\ \{123, 132, 213, 231, 312, 321\} = \Omega & for \quad x > 6 \end{cases}$$

حيث نعلم أنَّ كلاًّ من $\varnothing$ و Ω هي عناصر من 2^{Ω}.

إن هذا المتغيِّر العشوائيّ هو أبسط أنواع المتغيِّرات العشوائيّة على الإطلاق، ويُقال عنه إنَّه خاضع للتوزيع وحيد النقطة بـ مَعْلَمة Parameter $c = 6$. أي أنَّ لتوزيع هذا المتغيِّر العشوائيّ مَعْلَمة وحيدة هي القيمة c التي يأخذها هذا المتغيِّر العشوائيّ، ونلاحظ هنا أنَّ:

$$\mathbf{P}(X = c) = \mathbf{P}\big(\{\omega \in \Omega \mid X(\omega) = c\}\big) = \mathbf{P}(\Omega) = 1$$

٢- ليكن $\Omega = \{HH, HT, TH, TT\}$ فضاء الحوادث الابتدائيّة لتجربة رمي قطعة نقود لمرَّتين متتاليتين، و لنأخذ X تطبيقاً حقيقياً معرّفاً على Ω من خلال العلاقة الآتية:

$$X(\omega) = \begin{cases} -1 & for \quad \omega = HH, HT, TH \\ 2 & for \quad \omega = TT \end{cases}$$

فنجد أنَّ:

$$\{\omega \in \Omega \mid X(\omega) < x\} = \begin{cases} \varnothing & for \quad x \le -1 \\ \{HH, HT, TH\} & for \quad -1 < x \le 2 \\ \{HH, HT, TH, TT\} = \Omega & for \quad x > 2 \end{cases}$$

ولكن نعلم أنَّ $\{HH, HT, TH\}$، $\varnothing$ و Ω هي عناصر من 2^{Ω}، وهذا يعني أنَّ التطبيق X هو متغيِّر عشوائيّ على Ω، ويُقال عن هذا المتغيِّر العشوائيّ (الذي يمكن له أن يأخذ إحدى قيمتين فقط) إنَّه خاضع للتوزيع ثنائيّ النقطة في $x_1 = 2$ و $x_2 = -1$ بمَعْلَمَة p. إنّ قيمة المَعْلَمة p هي قيمة الاحتمال $\mathbf{P}(X = x_1)$ والذي يُدعى احتمال النجاح، وفي الحالة الخاصّة عندما يصبح لدينا $x_1 = 1$ و $x_2 = 0$ فإنَّه يُقال عن هذا المتغيِّر العشوائيّ إنَّه برنوليّ، وذلك نسبة إلى الرياضياتي السويسري يعقوب برنولي Jacob Bernoulli (1655-1705) الذي كان أول من قدَّم دراسة رياضياتية متقنة لهذه المسألة.

▶

٦-١-٧- ملاحظة

قياساً على المثال الأخير في الفقرة السابقة يُقال عن كلّ تجربة عشوائيّة تتمخَّض عنها إحدى نتيجتين فقط إنَّها تجربة برنوليّة، فعلى سبيل جميع التجارب الآتية تُصنَّف تجارب برنوليّة:

- رمي قطعة نقود لمرَّة واحدة (فتكون النتيجة: صورة أو شعار).
- إطلاق رمية على هدف من قبل شخص ما (فتكون النتيجة: إصابة أو عدم الإصابة).
- تقدّم طالب إلى اختبار (فتكون النتيجة: نجاح أو رسوب).
- تطبيق عقار دوائي على مريض (فتكون النتيجة: الدوء فعّال أو غير فعّال).
- تحليل عقار دوائي (فتكون النتيجة: مطابق للمواصفات أو غير مطابق للمواصفات).

وهكذا دواليك.

٦ - ٢
دالَّة التوزيع لمتغيِّر عشوائي

في كثير من الحالات، وخاصة لدى الجوانب التطبيقية والعمليّة، لا يظهر فيها المتغيِّر العشوائيّ بشكله الصريح، ولذا يستعاض عنه بتقديم ما يُعْرف باسم **دالّة توزيع** Distribution Function المتغيِّر العشوائيّ، ويطلق عليها البعض اسم **دالّة التوزيع التراكمية** Cumulative Distribution Function أيضاً، وهذه الدالّة تساعدنا في حساب احتمالات متعلِّقة بمتغيِّر عشوائيّ. إنَّ التعريف لهذا المفهوم تقدِّمه لنا الفقرة الآتية.

٦-٢-١- تعريف (دالّة توزيع متغيِّر عشوائيّ)

ليكن X متغيِّراً عشوائياً على فضاء حوادث ابتدائيّة Ω، ولنعرّف من أجل هذا المتغيِّر العشوائيّ دالّة حقيقيّة $\mathbf{F}_X$ على $\mathbb{R}$ من خلال العلاقة الآتية:

$$\mathbf{F}_X : \mathbb{R} \longrightarrow \mathbb{R}$$

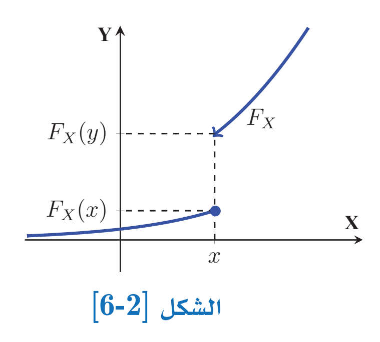
$$x \mapsto \mathbf{F}_X(x) := \mathbf{P}\Big(\{\omega \in \Omega \,|\, X(\omega) < x\}\Big) \qquad [6\text{–}3]$$

عندئذٍ تُدعى $\mathbf{F}_X$ **دالّة التوزيع الاحتماليّة** للمتغيِّر العشوائيّ X، وعلى سبيل التبسيط والاختصار يُقال **دالّة التوزيع** للمتغيِّر العشوائيّ X.

٦-٢-٢- ملاحظات

١- كما هو واضح من هذا التعريف فإنَّ المجال المقابل (**مجموعة القيم**) لدالّة التوزيع $\mathbf{F}_X$ هي الفترة $[0\,,\,1]$.

٢- إذا حصل لدالَّة التوزيع $\mathbf{F}_X$ انقطاع في موضعٍ x من $\mathbb{R}$ فإنَّها تقوم بقفزة نحو الأعلى عند هذه النقطة (انظر الشكل [6-2])، ونشير هنا إلى أنَّ الدائرة المغلقة على الرسم البياني للدالَّة $\mathbf{F}_X$ عند النقطة x تدلّ على قيمة الدالّة $\mathbf{F}_X$ عند هذه النقطة، بينما يدلّ رأس السهم في الرسم البياني للدالَّة $\mathbf{F}_X$ على قيمة الدالّة $\mathbf{F}_X$ عند أول قيمة $x < y$.



الشكل [6-2]

٣- إذا كانت $\mathbf{F}_X$ دالّة توزيع متغيِّر عشوائيّ X معرَّف على فضاء حوادث ابتدائيّة Ω، فعندئذٍ من أجل $-\infty < a < b < +\infty$ تكون العلاقة الآتية مُحقَّقة:

$$\mathbf{P}(a \leq X < b) = \mathbf{P}(X < b) - \mathbf{P}(X < a) = \mathbf{F}_X(b) - \mathbf{F}_X(a) \qquad [6\text{–}4]$$

وهذه العلاقة تساعدنا في حساب احتمالات متعلِّقة بالمتغيِّر العشوائيّ X إذا كانت دالّة توزيعه معلومة.

◄ ٦-٢-٣- أمثلة

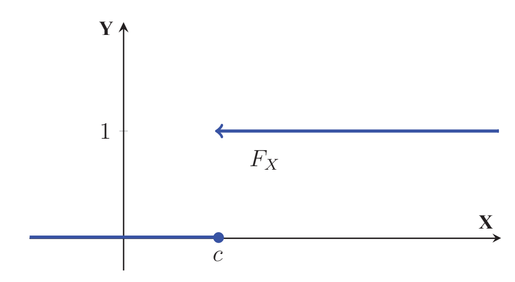
١- بالعودة إلى المثال (١) من (٦-١-٦) حيث لدينا X متغيِّر عشوائيّ معرَّف على Ω من خلال العلاقة:

$$X(\omega) = c$$

مع c ثابت عـددي، فعندئذٍ نجد أنّ دالّة توزيع هذا المتغيِّر العشوائيّ هي:

$$\mathbf{F}_X(x) = \mathbf{P}(X < x) = \begin{cases} 0 & \text{for } x \leq c \\ 1 & \text{for } x > c \end{cases}$$

وذلك لأنّه لدينا $\mathbf{P}(\varnothing) = 0$ و $\mathbf{P}(\Omega) = 1$ ، وأمّا رسمها البياني فله العرض الآتي:



الشكل [6–3] (الرسم البياني لدالّة توزيع المتغيِّر العشوائيّ X)

حيث نلاحظ أنّ لها قفزةً واحدةً نحو الأعلى عند النقطة c، وكذلك يُلاحظ أنّ لهذه الدالّة شكل درجة السلالم، ولذلك يُقال عن هذا النوع من الدوال إنّها دالّة درجية، وبما أنّ لها قفزةً واحدةً فقط عند c فهي أبسط دالّة درجية على الاطلاق (درجة واحدة فقط).

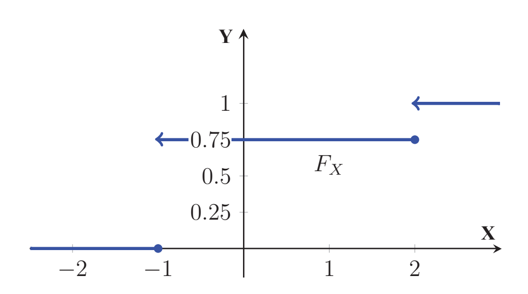
٢- بالرجوع إلى المثال (٢) من (٦-١-٦) ، فإذا افترضنا أنّ قطعة النقود متوازنة، وأنّ الحصول على شعارين هو النجاح (أي أن يأخذ X القيمة 2)، فإنّه سيكون لدينا بسبب توازن قطعة النقود ما يلي:

$$p = \mathbf{P}(X = 2) = \mathbf{P}\left(\{TT\}\right) = \frac{1}{4}$$

ومن ثمَّ نجد لدالَّة توزيـع هذا المتغيِّر العشوائيّ العرض الآتي:

$$\mathbf{F}_X(x) = \begin{cases} 0 & \text{for } x \leq -1 \\ \dfrac{3}{4} & \text{for } -1 < x \leq 2 \\ 1 & \text{for } x > 2 \end{cases}$$

وذلك لأنَّه لدينا $\mathbf{P}(\varnothing) = 0$ و $\mathbf{P}\left(\{HH, HT, TH\}\right) = \frac{3}{4}$ و $\mathbf{P}(\Omega) = 1$، وبالتالي يكون لدالَّة توزيع هذا المتغيِّر العشوائيّ العرض البياني الآتي:



الشكل [6–4] (الرسم البياني لدالَّة توزيع المتغيِّر العشوائيّ X)

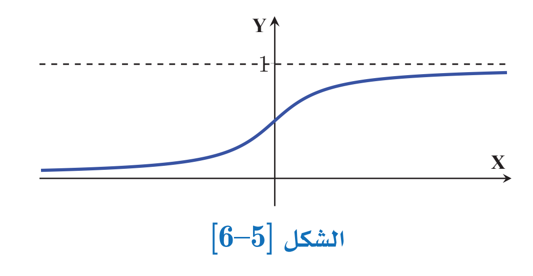
حيث نلاحظ أنَّ لهذه الدالّة شكل دالّة درجية بقفزتين (درجتين) أيضاً، ولو أردنا (على سبيل المثال) حساب $\mathbf{P}(-2 \leq X < 1)$، فإنَّنا نجد باستخدام العلاقة [6–4] ما يلي:

$$\mathbf{P}(-2 \leq X < 1) = \mathbf{F}_X(1) - \mathbf{F}_X(-2) = 0.75 - 0 = 0.75$$

▶

٦-٢-٤- ملاحظة

توجد متغيِّرات عشوائيّة لها دوال توزيع مستمرة (أو متصلة، أي الخط البياني لدالّة التوزيع مستمر ولا انقطاع فيه) على $\mathbb{R}$ بأكملها كما في الشكل الآتي:



الشكل [6–5]

وبناءً على ذلك يوجد نوعين على الأقل من المتغيرات العشوائيّة. أحدها له دوال توزيع درجية، والآخر له دوال توزيع مستمرة، ومن ثمَّ يمكننا أن نذكر تصنيفين على الأقل للمتغيِّرات العشوائيّة وهما:

- المتغيِّرات العشوائيّة المتقطِّعة (الرسم البياني لدوال توزيعها على شكل درجات السلالم).
- المتغيِّرات العشوائيّة المستمرَّة (الرسم البياني لدوال توزيعها خطوط متصلة وليس فيها قفزات).

وسنقوم في الفقرة التالية بتقديم دراسة مبسّطة عن كلّ نوعٍ من هذين النوعين، وسنبدأها بالمتغيِّرات العشوائيّة المتقطِّعة.

٦ - ٣
المتغيِّرات العشوائية المتقطِّعة

سنبدأ دراسة المتغيِّر العشوائيّ المتقطِّع بتقديم تعريفه ودالّة توزيعه على النحو الآتي.

٦-٣-١- تعريف (المتغيِّر العشوائيّ المتقطِّع)

ليكن X متغيِّراً عشوائياً على فضاء حوادث ابتدائيّة Ω، فإذا كانت مجموعة قيم X منتهيةً أو غير منتهيةٍ ولكن قابلةً للعدِّ، فعندئذٍ يُقال عن المتغيِّر العشوائيّ X إنَّه متقطِّع.

٦-٣-٢- أمثلة

١- ترد طلبيات إلى موظف بشكل مستقل كلّ منها عن الأخرى، فإذا كانت الطلبية مستكملة للثبوتيات المطلوبة فإنَّه يستقبل الطلبية التي تليها، وهكذا على هذا النحو إلى أن تصله طلبية ليست مستكملة الثبوتيات فيتوقف عن استلام أية طلبية أخرى في ذلك اليوم. فإذا علمت أنَّه يمكن أن يرد في اليوم الواحد 100 طلبية على الأكثر، وأنَّ احتمال أن تكون أية طلبية مستكملة للثبوتيات يساوي 0.95، فما هو احتمال أن يمتنع الموظف عن استلام الطلبية الثامنة؟

الحل: لنفترض أنَّ X متغيِّر عشوائيّ يرصد عدد الطلبيات المستلمة حتى الحصول على أول طلبية غير مستكملة الثبوتيات، فعندئذٍ نجد أنَّ هذا المتغيِّر العشوائيّ متقطِّع لأنَّ عدد القيم التي يمكن له أن يأخذها منتهٍ (القيم هي 1، 2 و... و100)، ومن ثمَّ يمكننا أن نعبِّر عن الاحتمال المطلوب من خلال الاحتمال $\mathbf{P}(X=7)$. من جهة أخرى نلاحظ أنَّ كلّ عملية تدقيق للطلبية هي تجربة برنوليّة (مستكملة الثبوتيات أو غير مستكملة الثبوتيات) فيها احتمال النجاح $p=0.05$. في هذه المسألة لدينا النجاح هو الحصول على طلبية غير مستكملة الثبوتيات لأنَّها هي التي ستقرِّر التوقف عن الاستلام، ولكن هذا ليس شرطاً ملزماً، فمن الممكن أن نأخذ الحصول على طلبية غير مستكملة الثبوتيات هو النجاح أيضاً. كذلك نلاحظ أنَّ هذه التجارب تتمّ بشكل مستقل كلّ منها عن الآخرى (لأنَّه لا يعلم شيئ عن طبيعة الطلبية المقدَّمة)، ومن ثمَّ يكون احتمال التوقف عن استلام الطلبية الثامنة هو احتمال وصول ست طلبيات مستكملة الثبوتيات والسّابعة غير مستكملة الثبوتيات، ومن ثمَّ بحسب قاعدة الضرب (في العدّ) يكون الاحتمال المطلوب هو:

$$(1-p)\cdot(1-p)\cdot(1-p)\cdot(1-p)\cdot(1-p)\cdot(1-p)\cdot p$$

وهي حالة في الحساب لا ثانية لها (أي لا تتمّ إلاّ بهذا التسلسل)، وبالتالي يمكننا أن نكتب:

$$\mathbf{P}(X=7)=(1-p)\cdot(1-p)\cdot(1-p)\cdot(1-p)\cdot(1-p)\cdot(1-p)\cdot p=\cdot(1-p)^6\cdot p$$
$$=(0.05)^6\cdot(0.95)=0.000000015$$

إنَّ المتغيِّر العشوائيّ المتقطّع الذي يتميَّز بالعلاقة:

$$\mathbf{P}\left(X=k\right)=(1-p)^{k-1}\cdot p \qquad ; k=1,2,3,\ldots$$

يُدعى **متغيِّرًا عشوائياً هندسياً** Geometric Random Variable بمَعْلَمة p ، علماً أنَّ $0<p<1$، وأنَّ القيم التي يأخذها هذا المتغيِّر العشوائيّ هي قيم صحيحة موجبة تماماً وعددها قد يكون غير منتهٍ ولكن قابل للعدّ على الأكثر، وبناءً على هذا فإنَّ المتغيِّر العشوائيّ X في مثالنا هذا هو من المتغيِّرات العشوائيّة الشهيرة، وله توزيع هندسي بمعْلَمة $p=0.05$.

٢- يوجد 20 قضية في إحدى المحاكم منها 7 قضايا اقتصادية (قضايا تتعلَّق بالاقتصاد). قام أحد القضاة بسحب عشوائيّ لخمس قضايا منها دفعة واحدة من أجل دراستها، فإذا علمت أنَّ لجميع القضايا النصيب نفسه في السحب (أو الاختيار)، فما هو احتمال أن يكون لدى هذا القاضي ثلاث قضايا اقتصادية؟

الحل: لنفترض أنَّ X متغيِّر عشوائيّ راصد لعدد قضايا الاقتصاد في القضايا التي سُحِبت، فعندئذٍ نجد أنَّ هذا المتغيِّر العشوائيّ متقطِّع لأنَّ القيم التي يمكن له أن يأخذها ست فقط (وهي 0، 1، 2، 3، 4 و5)، ومن ثمَّ يمكننا أن نعبِّر عن الاحتمال المطلوب من خلال الاحتمال $\mathbf{P}\left(X=3\right)$.

من جهة أخرى نلاحظ هنا أنَّ عمليات السحب ليست تجارب برنوليّة (لأنَّ السحب يتمّ على دفعة واحدة)، ولكن نلاحظ إمكانية تطبيق مبدأ لابلاس في الاحتمالات من أجل حساب الاحتمال المطلوب، حيث لدينا عدد الحالات الملائمة لحادث وجود ثلاث قضايا اقتصادية من القضايا الخمس التي سُحِبت يساوي $(7C3)\cdot(13C2)$، وأمّا عدد الحالات الممكنة للتجربة فإنَّه يساوي $(20C5)$، ومن ثمَّ يكون (بحسب مبدأ لابلاس في الاحتمالات) الاحتمال المطلوب هو:

$$\mathbf{P}\left(X=3\right)=\frac{(7C3)\cdot(13C2)}{(20C5)}=\frac{\binom{7}{3}\cdot\binom{13}{2}}{\binom{20}{5}}=\frac{\binom{7}{3}\cdot\binom{20-7}{5-3}}{\binom{20}{5}}=\frac{(35)\cdot(78)}{15504}=0.176$$

إنَّ المتغيِّر العشوائيّ الذي يتميَّز بالعلاقة الآتية:

$$\mathbf{P}\left(X=k\right)=\frac{\binom{M}{k}\cdot\binom{N-M}{n-k}}{\binom{N}{n}} \qquad ; k=0,1,2,\cdots,\min\{n,M\}$$

يُدعى **متغيِّرًا عشوائياً فوق هندسي** Hypergeometric Random Variable بمَعَالم N وM وn ، علماً أنَّ N وM وn أعداد طبيعية محدودة مع $1\le n\le N$ و $1\le M<N$، والقيم x_i التي يأخذها X هي قيم صحيحة غير سالبة، وهذا يعني أنَّ المتغيِّر العشوائيّ X في مثالنا هذا هو من المتغيِّرات العشوائيّة الشهيرة أيضاً، وله توزيع فوق هندسي بمَعَالم $N=20$ ، $M=7$ و $n=5$.

٣- يقوم شخص بالرمي على هدف دون أي يعلم نتيجة الإصابة، فإذا علمت أنَّ احتمال إصابته للهدف في أيّة رمية تساوي 0.85، وأنَّه أطلق على الهدف 3 رميات، فما هو احتمال أن يكون قد أصاب برميتين منها؟

✍ **الحل:** لنفترض أنَّ X هو متغيِّر عشوائيّ يرصد عدد الإصابات خلال الرميات الثلاث، فعندئذٍ نجد أنَّ هذا المتغيِّر العشوائيّ متقطِّع لأنَّ القيم التي يمكن له أن يأخذها أربع فقط (وهي 0، 1، 2 و3)، ومن ثمَّ يمكننا أن نعبِّر عن الاحتمال المطلوب من خلال الاحتمال $\mathbf{P}(X=2)$. من جهة أخرى نلاحظ أنَّ كلّ عملية إطلاق على الهدف هي تجربة برنوليّة (يصيب أو لا يصيب) فيها احتمال النجاح $p=0.85$، وأنَّ هذه التجارب تتمّ بشكل مستقل كلّ منها عن الأخرى (لأنَّه لا يعلم شيء عن نتيجة الإصابة للهدف)، ومن ثمَّ بحسب قاعدة الضرب (في العدّ) يكون لاحتمال الإصابة برميتين من أصل ثلاث أطلقت على الهدف لها أحد العروض الآتي:

$$p \cdot p \cdot (1-p) \quad \text{أو} \quad p \cdot (1-p) \cdot p \quad \text{أو} \quad (1-p) \cdot p \cdot p$$

وبما أنَّ حوادث إصابة الهدف وفقاً لهذه التسلسل من النتائج هي حوادث متنافية مثنى مثنى (كما هو واضح) فإنَّ الاحتمال سيكون مساوياً للمقدار الآتي:

$$p \cdot p \cdot (1-p) + p \cdot (1-p) \cdot p + (1-p) \cdot p \cdot p$$

ومن ثمَّ يمكننا أن نكتب:

$$\mathbf{P}(X=2) = p \cdot p \cdot (1-p) + p \cdot (1-p) \cdot p + (1-p) \cdot p \cdot p = 3\,[p^2 \cdot (1-p)]$$

لكن بملاحظة أنَّ العدد 3 هو في الواقع التوافيق لثلاثة أشياء أُختير منها شيئين بآنٍ واحدٍ، وأنَّ أس المقدار $(1-p)$ يمكن كتابته من خلال $(3-2)$ (عدد المحاولات مطروحاً منه عدد النجاحات) فإنَّه يمكننا أن نكتب المقدار السابق على النحو الآتي:

$$\mathbf{P}(X=2) = \binom{3}{2} p^2 \cdot (1-p) = \binom{3}{2} (0.85)^2 \cdot (0.15) = 0.3825$$

وبالتالي نجد أنَّ الاحتمال المطلوب هو $\mathbf{P}(X=2) = 0.3825$.

سنلاحظ فيما بعد أنَّ المتغيِّر العشوائيّ X في مثالنا هذا هو من المتغيِّرات العشوائيّة الشهيرة أيضاً، وأنَّه سيكون خاضعاً للتوزيع الحدّاني بمَعْلَمتين $n=3$ و $p=0.85$.

▶

٦-٣-٣- ملاحظات

١- يمكن تمييز المتغيِّر العشوائيّ المتقطِّع من خلال الاحتمالات $\mathbf{P}(X=x_i)$ من أجل كلّ القيم الممكنة لـ i، وتُدعى هذه القيم للاحتمالات بـ **الكتل الاحتماليّة** Probability Masses للمتغيِّر العشوائيّ X، ويُرمز لها عادةً بـ p_i، أي أنَّه يُكْتَب:

$$p_i = \mathbf{P}(X=x_i) \qquad ; \forall i$$

٢- إذا كان فضاء الحوادث الابتدائيّة Ω منتهٍ، فعندئذٍ ستكون مجموعة قيم المتغيِّر العشوائيّ المتقطِّع X منتهية أيضاً، وفي هذه الحالة يُقال عن المتغيِّر العشوائيّ X إنَّه بسيط.

٣- على سبيل التبسيط والتوضيح سوف نخصّص دراستنا على متغيِّرات عشوائيّة لها عدد منتهٍ من القيم، أي أنَّنا سنأخذ مجموعة قيم المتغيِّر العشوائيّ المتقطِّع X من الشكل:

$$\mathbf{X} = \{x_i \mid i = 1, 2, \dots, k\}$$

٦-٣-٤- التمثيل الجدولي والبياني للمتغيِّرات العشوائيّة المتقطِّعة

انطلاقاً من مُعرفَّة القيم x_i و p_i من أجل كلّ القيم المكنة لـ i يمكننا تقديم المتغيِّر العشوائيّ البسيط X (ولو نظرياً على الأقل) وفق طرائق أخرى منها على سبيل المثال:

أ- التمثيل الجدولي:

بفرض أنَّ X متغيِّر عشوائيّ بسيط على فضاء حوادث ابتدائيّة Ω وبمجموعة قيم:

$$\mathbf{X} = \{x_i \mid i = 1, 2, \dots, k\}$$

فعندئذٍ يمكن تقديم المتغيِّر العشوائيّ X جدولياً من خلال جدول بصفين يدوَّن في صفه العلوي القيم x_i للمتغيِّر العشوائيّ X، وأمَّا في صفه السفلي فيتمّ تدوين الكتل الاحتماليّة p_i المقابلة لتلك القيم x_i على النحو الموضَّح في الجدول الآتي:

الجدول [6–1]

x_i قيم X	x_1	x_2	...	x_k
$p_i = \mathbf{P}(X = x_i)$	p_1	p_2	...	p_k

إنَّ هذا الجدول يُدعى **جدول توزيع** المتغيِّر العشوائيّ X أيضاً، وهذه التسمية خاصّة بالمتغيِّرات العشوائيّة المتقطِّعة البسيطة فقط. كذلك يجب الأخذ بالحسبان أن تكون العلاقة $\sum_{i=1}^{k} p_i = 1$ مُحقَّقة.

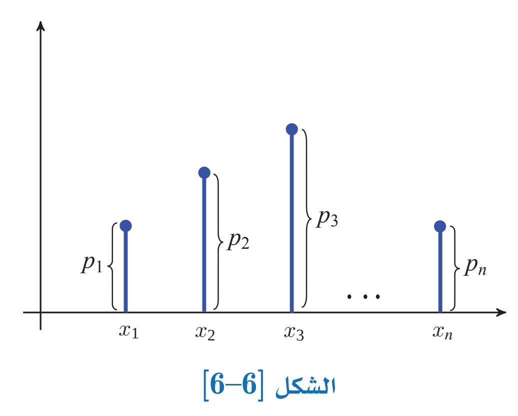
ب- التمثيل البياني:

بفرض أنَّ X متغيِّر عشوائيّ بسيطٌ على فضاء حوادث ابتدائيّة Ω وبمجموعة قيم:

$$\mathbf{X} = \{x_i \mid i = 1, 2, \dots, k\}$$

فعندئذٍ يمكن أن يُمثَّل هذا المتغيِّر العشوائيّ بيانياً من خلال رسم محورين إحداثيين متعامدين، ومن ثمَّ رسم عمود (يُمثِّل قفزة المتغيِّر العشوائيّ) بارتفاع يساوي p_i فوق القيمة x_i (التي تُدعى موضع القفزة للمتغيِّر العشوائيّ) وذلك من أجل كلّ القيم ممكنة لـــ i.

نشير هنا إلى أنَّ بعض المراجع تضع في نهاية كلّ عمود دائرة صغيرة مصمتة كتعبير عن موضع الكتلة الاحتماليّة للمتغيِّر العشوائيّ X عند موضع القفزة، وأمّا ارتفاع العمود (وهو الأهمّ في الرسم) فإنَّه يعبِّر عن قيمة الكتلة الاحتماليّة للمتغيِّر العشوائيّ X عند موضع القفزة. وهكذا يمكن للمتغيِّر العشوائيّ X أن يُمثَّل بيانياً من خلال الشكل الآتي:



الشكل [6–6]

٦-٣-٥- ملاحظات

١- في الأشكال القادمة الممثِّلة للمتغيِّرات العشوائيّة المتقطِّعة سوف نتجاوز عن رسم الدوائر الصغيرة المصمتة في نهايات الأعمدة وذلك لأنَّ المهمّ في هذه العروض هو ارتفاعات الأعمدة عند مواضع القفزات.

٢- يُلاحظ عدم جدوى تقديم المتغيِّرات العشوائيّة البسيطة بالطريقة الجدوليّة أو البيانيّة عندما تكون مجموعة قيمها كبيرة، وأمّا إذا كانت القيم التي تأخذها هذه المتغيِّرات العشوائيّة قليلة نسبياً فإنَّ هاتين الطريقتين تعطيان عروضاً مقبولة.

٣- تجدر الإشارة هنا إلى أنَّ التوزيعات النّاتجة عن متغيِّرات عشوائيّة متقطِّعة تُدعى **توزيعات متقطِّعة**.

٦-٣-٦- أمثلة

١- لنفترض أنَّ متغيِّراً عشوائياً X قُدِّم من خلال الجدول الآتي:

الجدول [6–2]

i	1	2	3	4	5	6	7	8	9
قيم X x_i	0	1	2	3	4	5	6	7	8
p_i	0.0038	0.0313	0.1094	0.2188	0.2734	0.2188	0.1094	0.0313	0.0038

فنجد أنَّ هذا المتغيِّر العشوائيّ متقطِّعٌ بسيطٌ، ولكن يجب التحقّق أولاً من صحة العلاقة $\sum_i p_i = 1$، حيث لدينا:

$$\sum_{i=1}^{9} p_i = 0.0038 + 0.0313 + 0.1094 + 0.2188 + 0.2734 + 0.2188 + 0.1094 + 0.0313 + 0.0038 = 1$$

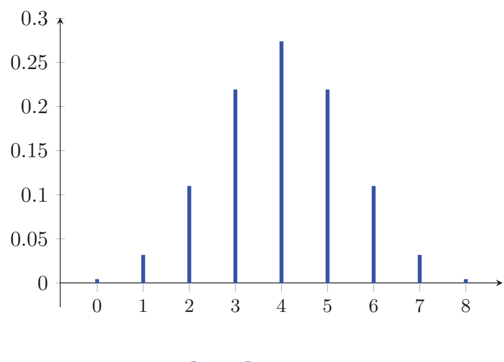
ومن ثمَّ العلاقات التي تميِّز هذا المتغيِّر العشوائيّ هي:

$\mathbf{P}(X=0)=0.0038$ & $\mathbf{P}(X=1)=0.0313$ & $\mathbf{P}(X=2)=0.1094$

$\mathbf{P}(X=3)=0.2188$ & $\mathbf{P}(X=4)=0.2734$ & $\mathbf{P}(X=5)=0.2188$

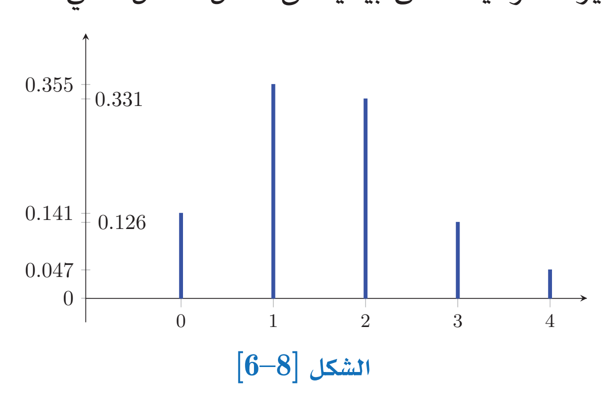
$\mathbf{P}(X=6)=0.1094$ & $\mathbf{P}(X=7)=0.0313$ & $\mathbf{P}(X=8)=0.0038$

وتمثيله البياني يقدِّمه لنا الشكل الآتي.



الشكل [6–7]

٢- ليكن لدينا X متغيِّراً عشوائياً مُعطى بيانياً من خلال الشكل الآتي:



الشكل [6–8]

فنجد لتمثيله الجدولي العرض الآتي:

الجدول [6–3]

i	1	2	3	4	5	المجموع
قيم X x_i	0	1	2	3	4	
p_i	0.141	0.355	0.331	0.126	0.047	1

ومن ثمَّ العلاقات التي تميِّز هذا المتغيِّر العشوائيّ المتقطِّع البسيط هي:

$$\mathbf{P}(X=0)=0.141 \quad \& \quad \mathbf{P}(X=1)=0.355 \quad \& \quad \mathbf{P}(X=2)=0.331$$

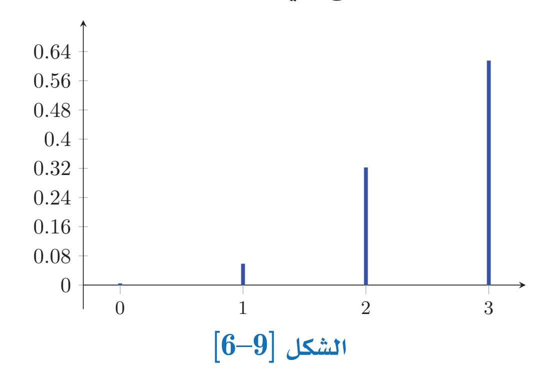
$$\mathbf{P}(X=3)=0.126 \quad \& \quad \mathbf{P}(X=4)=0.047$$

٣- بالعودة إلى المثال (٣) من (٦-٣-٢) يمكننا عرض ذلك المتغيّر جدولياً كما يلي:

الجدول [6–4]

i	1	2	3	4	المجموع
x_i قيم X	0	1	2	3	
p_i	0.003375	0.057375	0.325125	0.614125	**1**

حيث لدينا $n=3$ هو عدد التجارب البرنوليّة المستقلّة التي تمّ تنفيذها، x_i عدد النجاحات التي تحقّقت خلال الـ n تجربة، $p=0.85$ هو احتمال النجاح في أية تجربة، وأمّا تمثيله البياني فله الشكل الآتي.



الشكل [6–9]

٦-٣-٦- ملاحظات

١- يُلاحظ ممّا سبق أنّه من أجل متغيِّر عشوائيّ متقطّع بسيط وبقيم قليلة العدد يمكننا استخدام أيّ من التمثيلات التي ذكرناها سابقاً لاستنتاج بقية التمثيلات الأخرى للمتغيِّر العشوائيّ.

٢- لقد لاحظنا في المثال (٣) من (٦-٣-٢) أنّه إذا كان لدينا تجربة برنوليّة باحتمال نجاح p ، وقمنا بتكرار هذه التجربة لـ n مرَّة متتالية وبشكلٍ مستقل كلّ منها عن الأخرى، وبفرض أنّ X متغيّر عشوائيّ راصد لعدد النجاحات (وليكن k) خلال التجارب المتتالية التي نفِّذت، فإنّ العلاقة المميِّزة لهذا المتغيِّر العشوائيّ سيكون لها العرض الآتي:

$$\mathbf{P}(X=k)=\binom{n}{k}(p)^{k}\cdot(1-p)^{n-k} \qquad ; k=0,1,2,\ldots,n$$

حيث نلاحظ أنَّ القيم التي يأخذها X هي قيم صحيحة غير سالبة (عددها $n+1$ قيمة). إنَّ المتغيِّر العشوائيّ الذي يتميَّز بهذه العلاقة يُدعى متغيِّراً عشوائياً حِدَّانياً Binomial Random Variable، أو يُقال إنَّ له توزيع حِدَّاني Binomial Distribution بمَعْلَمتين n وp علماً أنَّ n عدد طبيعي مثبَّت و $0<p<1$.

٦-٣-٧- تعريف (المتغيرِّ العشوائيّ الحِدَّاني)

ليكن X متغيرِّاً عشوائياً على فضاء حوادث ابتدائيّة Ω وبمجموعة قيم $\mathbf{X}=\{0,1,2,...,n\}$ مع n عدد طبيعي مثبَّت. عندئذٍ يكون لهذا المتغيرِّ العشوائيّ توزيع حِدَّاني (أو ذي الحَدين) بمَعْلَمتين n وp مع $0<p<1$، إذا كان للعلاقة التي تميّزه العرض الآتي:

$$\mathbf{P}(X=k)=\binom{n}{k}\cdot p^k\cdot(1-p)^{n-k} \quad ;\ k=0,1,2,...,n \qquad [6\text{–}5]$$

إذاً. نستنتج ممّا سبق أنَّ المتغيرِّ العشوائيّ الحِدَّاني يرصد عدد النجاحات عند تنفيذ عدد منتهٍ من التكرارات لتجارب برنوليّة مستقلّة باحتمال نجاح معلوم.

◀ ٦-٣-٨- أمثلة

١- يوجد 1000 مصباح من نوع معيَّن في مستودع، منها 40 مصباحاً معيباً (غير مطابق للمواصفات). نقوم بسحب 10 مصابيح عشوائياً من المستودع ومن ثمَّ فحصها على التتالي بشكل مستقل كلّ منها عن الآخر (واحداً تلو الآخر)، فما هو احتمال:

أ- حصولنا على 8 مصابيح سليمة؟

ب- حصولنا على 8 مصابيح سليمة على الأقل؟

ج- عدم حصولنا على أي مصباح سليم؟

الحل: بملاحظة أنَّ كلّ عملية فحص لمصباح هي تجربة برنوليّة (سليم أو معيب) فيها احتمال النجاح:

$$p=\frac{1000-40}{1000}=\frac{960}{1000}=0.96$$

وأنَّ هذه التجارب تتمّ بشكل مستقل كلّ منها عن الأخرى (لأنّنا لا نعلم شيء عن نتيجة الفحص مسبقاً)، فعندئذٍ بفرض أنَّ X هو متغيرّ عشوائيّ راصد لعدد المصابيح السليمة خلال الفحوصات العشر، فإنَّه سيكون لهذا المتغيرِّ العشوائيّ توزيع حِدَّاني بمَعْلَمتين $n=10$ و $p=0.96$، ومن ثمَّ يكون لدينا من أجل الطلب:

أ- الاحتمال المطلوب هو $\mathbf{P}(X=8)$ (حيث لدينا $k=8$)، وبحسب العلاقة $[6\text{–}5]$ نجد:

$$\mathbf{P}(X=8)=\binom{10}{8}p^8\cdot(1-p)^{10-8}=(45)\cdot(0.96)^8\cdot(0.04)^2=0.052$$

ب- الاحتمال المطلوب هو $\mathbf{P}(X\geq 8)$ والذي يكتب على النحو الآتي:

$$\mathbf{P}(X\geq 8)=\mathbf{P}(X=8)+\mathbf{P}(X=9)+\mathbf{P}(X=10)$$

وباستخدام العلاقة $[6\text{–}5]$ في حساب كلّ حدٍ من حدود الطرف الأيمن من العلاقة السابقة نجد أنَّ الاحتمال المطلوب هو:

$$\mathbf{P}(X\geq 8)=(0.05194)+(0.27701)+(0.66483)=0.99378$$

ج- لدينا الاحتمال المطلوب هو $\mathbf{P}(X=0)$، وباستخدام العلاقة [5–6] نجد أنَّ الاحتمال المطلوب هو:

$$\mathbf{P}(X=0)=\binom{10}{0}(0.96)^0\cdot(0.04)^{10-0}=1\times1\times0.00000000000001\approx0$$

٢- يقوم شخص ما بكتابة نص على الحاسب الآلي، وبحيث أنَّ كتابة كلّ حرف مستقلّة عن الأخرى، وبفرض أنَّ احتمال أن يخطئ في كتابة أي حرف يساوي 0.015، فعندئذٍ:

أ- إذا كان النص مكوَّناً من 1500 حرفاً، فما هو احتمال أن لا يخطئ أبداً في كتابة هذا النص؟

ب- إذا كان النص مكوَّناً من 15 حرفاً فقط، فما هو احتمال أن يخطئ في كتابة حرفين على الأقل؟

الحل: نلاحظ أنَّ كلّ كتابة كلّ حرف هي تجربة البرنوليّة (**إمّا أن يُخطأ في كتابته أو لا يخطئ**)، وأنَّ عملية الكتابة للنص هي تجارب برنوليّة مستقلّة بعضها عن البعض الآخر، ومن ثمَّ بفرض أنَّ X متغيِّر عشوائيّ راصد لعدد الأخطاء المرتكبة في كتابة النص، فعندئذٍ من أجل الطلب:

أ- سيكون للمتغيِّر العشوائيّ X توزيع حِدَّاني بمَعْلَمتين $n=1500$ و $p=0.015$، وبالتالي احتمال أن لا يخطئ أبداً في كتابة هذا النص يساوي:

$$\mathbf{P}(X=0)=\binom{1500}{0}(0.015)^0\cdot(1-0.015)^{1500-0}$$
$$=0.985^{1500}=0.00000000014\approx0$$

وهذه النتيجة الضئيلة جداً جداً لقيمة الاحتمال يُعبِّر عنها بالقول: إنَّ حادث أن لا يخطئ الكاتب في كتابة هذا النص أبداً هو **حادث شبه مستحيل**. ذلك أنَّ الحادث شبه المستحيل هو حادث A مع $A\neq\varnothing$ ومن أجله يكون $\mathbf{P}(A)=0$.

ب- سيكون للمتغيِّر العشوائيّ X توزيع حِدَّاني بمَعْلَمتين $n=15$ و $p=0.015$، ومن ثمَّ احتمال أن يخطئ الكاتب في كتابة حرفين على الأقل يساوي:

$$\mathbf{P}(X\geq2)=1-\left[\mathbf{P}(X=0)+\mathbf{P}(X=1)\right]$$

حيث لدينا:

$$\mathbf{P}(X=0)=\binom{15}{0}(0.015)^0\cdot(1-0.015)^{15-0}=(0.985)^{15}=0.797$$

$$\mathbf{P}(X=1)=\binom{15}{1}(0.015)^1\cdot(1-0.015)^{15-1}=15\times0.015\times0.8093=0.182$$

ومنه ينتج لدينا أنَّ الاحتمال المطلوب يساوي:

$$\mathbf{P}(X\geq2)=1-(0.797+0.182)=0.021$$

٣- بالعودة إلى المثال (٣) من (٦-٣-٢) نلاحظ أنَّه عدد التجارب البرنوليّة المستقلّة التي تمَّ تنفيذها يساوي 3، واحتمال النجاح في أية تجربة يساوي $p = 0.85$ ، ومن ثمَّ يكون للمتغيِّر العشوائيّ X (الراصد لعدد الإصابات خلال الرميات الثلاث) توزيع حِدَّاني بمَعْلَمتين $n = 3$ و $p = 0.85$.

٦-٣-٩- التوقع الرياضياتي (أو متوسط القيم) لمتغيِّر عشوائيّ متقطع

Mathematical Expectation of Discrate Random Variable

بالرجوع إلى الفصل الثاني من هذا الكتاب، والنظر في صيغة المتوسط الموزون لمجموعة من القيم x_i بأوزان w_i عندما تمسح i كلّ القيم الممكنة حيث لدينا:

$$\bar{x}_w = \frac{\sum_i x_i \cdot w_i}{\sum_i w_i}$$

وكذلك لو أمعنّا النظر في جداول التوزيعات الاحتماليّة للمتغيّرات العشوائيّة المتقطّعة التي مرَّت معنا سابقاً، فإنّنا سنلاحظ (وبوضوح تام) أنَّ قيم الاحتمالات $\mathbf{P}(X = x_i)$ تمثِّل في الواقع أوزاناً لـ x_i قيم المتغيّر العشوائيّ المتقطّع X ، ومن ثمَّ ستكون النسبة:

$$\frac{\sum_i x_i \cdot \mathbf{P}(X = x_i)}{\sum_i \mathbf{P}(X = x_i)}$$

هي المتوسط الموزون للقيم x_i عندما تمسح i كلّ القيم الممكنة لها، وبما أنَّ $\sum_i \mathbf{P}(X = x_i) = 1$ دوماً، فإنَّه سيصبح $\sum_i x_i \cdot \mathbf{P}(X = x_i)$ هو المتوسط الموزون للقيم x_i عندما تمسح i كلّ القيم الممكنة لها، والذي هو في الواقع متوسط قيم المتغيِّر العشوائيّ المتقطّع X. أمّا رياضياتياً فإنَّ المجموع السابق يُدعى التوقع الرياضياتي Mathematical Expectation للمتغيِّر العشوائيّ X ويُرمز له بـ $\mathbf{E}(X)$ ، وبهذا التمهيد يمكننا أن نقدِّم التعريف الآتي.

٦-٣-١٠- تعريف (التوقع الرياضياتي لمتغيِّر عشوائيّ متقطِّع)

ليكن X متغيِّراً عشوائياً متقطِّعاً على فضاء حوادث ابتدائيّة Ω وبمجموعة قيم:

$$\mathbf{X} = \{x_i \mid i = 1, 2, \ldots, k\}$$

فعندئذٍ يُعرَّف التوقع الرياضياتي لـ X من خلال العلاقة الآتية:

$$\mathbf{E}(X) := \sum_{i=1}^{k} x_i \, \mathbf{P}(X = x_i) \qquad [6\text{–}6]$$

٦-٣-١١- بعض خصائص التوقع الرياضياتي لمتغيرٍ عشوائيّ متقطع

١- إذا كانت قيم المتغيرّ العشوائيّ المتقطّع X تقع بين قيمتين t و u، فعندئذٍ سيكون لدينا:

$$t \le \mathbf{E}(X) \le u$$

وإذا أصبحت $t = u$ فعندئذٍ ينتج لدينا:

$$\mathbf{E}(X) = \mathbf{E}(t) = t \qquad [6–7]$$

أي أنَّ **التوقع الرياضياتي** للقيمة الثابتة هو القيمة نفسها، وكذلك ينتج من هذه الخاصّية أنَّه إذا كانت $t = 0$ (**أي أنَّ X متغيرٌ عشوائيّ غير سالب**)، فعندئذٍ سيكون $0 \le \mathbf{E}(X)$ أيضاً.

٢- إذا كان a وb عددين حقيقيين مع $a \ne 0$، فإنَّ العلاقات الآتية ستكون محقَّقة:

$$\mathbf{E}\left(a\,X + b\right) = a \cdot \mathbf{E}(X) + b \qquad [6–8]$$

$$\mathbf{E}\left(X - \mathbf{E}(X)\right) = 0 \qquad [6–9]$$

علماً أنَّ المتغيرِّ العشوائيّ $X - \mathbf{E}(X)$ يُدعى **متغيرّاً عشوائياً مركزياً**، ومن ثمَّ يكون التوقُّع الرياضياتي لأي متغيرِّ عشوائيّ مركزي يساوي الصفر دوماً.

٦-٣-١٢- أمثلة

١- الجدول الآتي يقدِّم لنا متغيرّاً عشوائياً من جدول توزيعه، والمطلوب حساب التوقع الرياضياتي لهذا المتغيرّ العشوائيّ.

الجدول [6–5]

قيم X x_i	1	2	3	4	المجموع
p_i	0.10	0.25	0.35	0.30	1

الحل: من أجل حساب التوقع الرياضياتي لهذا المتغيرّ العشوائيّ لدينا:

$$\mathbf{E}(X) = \sum_{i=1}^{4} x_i \cdot \mathbf{P}(X = x_i) = 1(0.10) + 2(0.25) + 3(0.35) + 4(0.30) = 2.85$$

٢- قامت إحدى مؤسسات بيع الأدوات المنزلية بتقديم عروض على سلعها بحيث يكون تخفيض السعر فيه على خمس أصناف 1 ، 2 ، 3 ، 4 و5 ، ثمَّ رصد شراء النّاس لهذه الأصناف حتى نفاذها، فكانت النتائج كما في الجدول الآتي:

الجدول [6–6]

مستوى التخفيض	5	4	3	2	1
الكمية المباعة	8	12	35	30	15
الأرباح بالريال	400	1140	2450	4050	1500

ولنقم بحساب متوسط كميّة المبيعات وكذلك الربح المتوقع للسلع المباعة.

الحل: من أجل حساب متوسط كميّة المبيعات سنفترض أنَّ X هو متغيِّر عشوائيّ راصد لكميّة المبيعات، فعندئذٍ سيكون لهذا المتغيِّر العشوائيّ القيم $x_1 = 15$ و $x_2 = 30$ و $x_3 = 35$ و $x_4 = 12$ و $x_5 = 8$ ، ومن ثمَّ يكون لدينا:

$$\mathbf{P}(X = 15) = \frac{15}{100} = 0.15 \quad \& \quad \mathbf{P}(X = 30) = \frac{30}{100} = 0.30$$

$$\mathbf{P}(X = 35) = \frac{35}{100} = 0.35 \quad \& \quad \mathbf{P}(X = 12) = \frac{12}{100} = 0.12$$

$$\mathbf{P}(X = 8) = \frac{8}{100} = 0.08$$

وبالتالي نجد أنَّ متوسط كميّة المبيعات للسّلع المُباعة يساوي:

$$\mathbf{E}(X) = \sum_{i=1}^{5} x_i \cdot \mathbf{P}(X = x_i) = (15 \times 0.15) + (30 \times 0.30) + (35 \times 0.35)$$
$$+ (12 \times 0.12) + (8 \times 0.08) = 25.58$$

وهذا يعني أنَّ متوسط كميّة المبيعات للسّلع المُباعة من كلّ صنف يساوي 26 قطعة تقريباً.

وأمّا من أجل حساب الربح المتوقع للسّلع المُباعة من كلّ صنف سنفترض أنَّ X متغيِّر عشوائيّ راصدٌ لأرباح الكميات المُباعة، فعندئذٍ سيكون لهذا المتغيِّر العشوائيّ القيم $x_1 = 1500$ و $x_2 = 4050$ و $x_3 = 2450$ و $x_4 = 1140$ و $x_5 = 400$، وبالتالي يكون لدينا:

$$\mathbf{P}(X = 1500) = \frac{1500}{7590} = 0.1976 \quad \& \quad \mathbf{P}(X = 4050) = \frac{4050}{7590} = 0.5336$$

$$\mathbf{P}(X = 2450) = \frac{2450}{7590} = 0.3228 \quad \& \quad \mathbf{P}(X = 1140) = \frac{1140}{7590} = 0.1502$$

$$\mathbf{P}(X = 400) = \frac{400}{7590} = 0.0527$$

ومن ثمَّ نجد أنَّ الربح المتوقع للسّلع المُباعة يساوي:

$$\mathbf{E}(X) = \sum_{i=1}^{5} x_i \cdot \mathbf{P}(X = x_i) = (1500 \times 0.1976) + (4050 \times 0.5336)$$
$$+ (2450 \times 0.3228) + (1140 \times 0.1502) + (400 \times 0.0527) = 3440.65$$

٣- بالعودة إلى المثال السابق (٢) من (٦-٣-٨)، فإذا علمنا أنَّ التوقع الرياضياتي لمتغيِّر عشوائيّ حِدَّاني X بمَعْلَمتين n وp يساوي $\mathbf{E}(X) = n \cdot p$ ، فما هو متوسط عدد الأحرف التي سيخطئ في كتابتها ذلك الشخص في كلا الحالتين (أ) و (ب)؟

الحل: من أجل الحالة:

أ- لدينا متوسط عدد الأحرف التي سيخطئ في كتابتها ذلك الشخص يساوي:

$$\mathbf{E}(X) = n \cdot p = 1500 \times 0.015 = 22.5 \approx 23$$ حرفاً

أي أنَّنا سنتوقع منه كتابة 23 حرفاً خاطئاً في هذه الحالة.

ب- لدينا متوسط عدد الأحرف التي سيخطئ في كتابتها ذلك الشخص يساوي:

$$\mathbf{E}(X) = n \cdot p = 15 \times 0.015 = 0.225 \approx 0$$

أي أنَّنا سنتوقع منه كتابة نصٍ خالٍ من الأخطاء في هذه الحالة.

▶

٦-٣-١٣- التّباين لمتغيرٍّ عشوائيّ متقطع

ليكن X متغيرّاً عشوائياً متقطّعاً على فضاء حوادث ابتدائيّة Ω وبمجموعة قيم:

$$\mathbf{X} = \left\{ x_i \mid i = 1, 2, \ldots, k \right\}$$

فعندئذٍ يُعرَّف التّباين لـ X (ويُرمز له $\mathbf{var}(X)$) من خلال العلاقة الآتية:

$$\mathbf{var}(X) = \mathbf{E}\left(X - \mathbf{E}(X)\right)^2 = \sum_{i=1}^{k} \left(x_i - \mathbf{E}(X)\right)^2 \cdot \mathbf{P}(X = x_i) \qquad [6\text{–}10]$$

٦-٣-١٤- ملاحظات

١- يمكن عرض العلاقة السابقة [6–10] على النحو الآتي:

$$\mathbf{var}(X) = \sum_{i=1}^{k} x_i^2 \cdot \mathbf{P}(X = x_i) - \left(\sum_{i=1}^{k} x_i \cdot \mathbf{P}(X = x_i)\right)^2 \qquad [6\text{–}11]$$

٢- يُطلق على الجذر التربيعي الموجب لتباين متغيرٍّ عشوائيّ X اسم "**الانحراف المعياري**" لـ X، ويرمز له عادة بـ σ. أي أنَّه لدينا:

$$\sigma = +\sqrt{\mathbf{var}(X)} \qquad [6\text{–}12]$$

ومن ثمَّ نلاحظ أنَّه يمكننا أن نكتب $\mathbf{var}(X) = \sigma^2$.

◀٦-٣-١٥- أمثلة

١- بالعودة إلى المثال (٢) من (٦-٣-٦) حيث لدينا متغيرٍّ عشوائيّ متقطِّع بجدول توزيع مُقدَّم كما يلي:

قيم X x_i	0	1	2	3	4	Total
p_i	0.141	0.355	0.331	0.126	0.047	1

ولنقم بحساب التوقع الرياضياتي، التّباين والانحراف المعياري لهذا المتغيرّ العشوائيّ.

الحل: من أجل حسابات التوقع الرياضياتي، التّباين والانحراف المعياري لمتغيرٍّ عشوائيّ متقطِّع له عدد قليل من القيم (نسبياً) يُفضَّل استخدام الجدول الآتي:

الجدول [6–7]

i	x_i	$\mathbf{P}(X = x_i)$	$x_i \cdot \mathbf{P}(X = x_i)$	$x_i^2 \cdot \mathbf{P}(X = x_i)$
1	0	0.141	0	0
2	1	0.355	0.355	0.355
3	2	0.331	0.662	1.324
4	3	0.123	0.369	1.107
5	4	0.047	0.235	1.175
Total	---	**1**	**1.621**	**3.961**

ومن ثمَّ يكون لدينا:

$$\mathbf{E}(X) = \sum_{i=1}^{5} x_i \cdot \mathbf{P}(X = x_i) = 1.621$$

$$\sum_{i=1}^{5} x_i^2 \cdot \mathbf{P}(X = x_i) = 3.961$$

وباستخدام العلاقة [6–11] نجد أنَّ تباين المتغيِّر العشوائيّ X يساوي:

$$\mathbf{var}(X) = \sum_{i=1}^{5} x_i^2 \cdot \mathbf{P}(X = x_i) - \left(\sum_{i=1}^{5} x_i \cdot \mathbf{P}(X = x_i) \right)^2 = 3.961 - (1.621)^2 = 1.333359$$

وأخيراً باستخدام العلاقة [6–12] نجد أنَّ الانحراف المعياري لهذا المتغيِّر العشوائيّ يساوي:

$$\sigma = +\sqrt{1.333359} = 1.1547$$

٢- بالعودة إلى المثال السابق (٢) من (٨-٣-٦)، فإذا عَلِمنا أنَّ التّباين لمتغيِّر عشوائيّ حِدَّاني X بمَعْلَمتين n وp يُعطى بالعلاقة الآتية:

$$\mathbf{var}(X) = n \cdot p \cdot (1 - p)$$

فما هي قيمة الانحراف المعياري للمتغيِّر العشوائيّ X في كلا الحالتين (أ) و (ب) ؟

الحل: لدينا من أجل الطلب:

أ- تباين المتغيِّر العشوائيّ X يساوي:

$$\mathbf{var}(X) = n \cdot p \cdot (1 - p) = 1500 \times 0.015 \times 0.985 = 22.1625$$

ومن ثمَّ تكون قيمة الانحراف المعياري للمتغيِّر العشوائيّ X هي:

$$\sigma = +\sqrt{\mathbf{var}(X)} = +\sqrt{22.1625} = 4.7077$$

أي أنَّ عدد الأحرف التي قد يخطئ في كتابتها قد تَصِل إلى 28 حرفاً، وإن قلَّت فقد تكون 18 حرفاً.

ب- تباين المتغيّر العشوائيّ X يساوي:

$$\textbf{var}(X) = n \cdot p \cdot (1-p) = 15 \times 0.015 \times 0.985 = 0.221625$$

ومن ثمَّ تكون قيمة الانحراف المعياري للمتغيّر العشوائيّ X هي:

$$\sigma = +\sqrt{\textbf{var}(X)} = +\sqrt{0.221625} = 0.47077$$

أي أنَّ عدد الأحرف التي قد يخطئ في كتابتها تكاد تكون معدومةً في هذه الحالة.

٦ - ٤
المتغيّرات العشوائية المستمرّة

لقد ذكرنا سابقاً أنَّ المتغيّرات العشوائيّة المستمرّة تتميَّز بأنَّ دوال توزيعاتها الاحتماليّة مستمرَّة على $\mathbb{R}$، ولكنَّ التعريف الدقيق لهذا النوع من المتغيّرات العشوائيّة لا يمكن تقديمه على مستوى هذا الكتاب، ولذلك سنكتفي بتقديم أحد أنواعه الشهيرة ودون الخوض في التفاصيل التي تقع خارج إطار هذا الكتاب.

٦-٤-١- التوزيع الطّبيعي Normal Distribution

يُعدّ التوزيع الطّبيعي من أكثر التوزيعات الاحتماليّة أهميةً وفائدةً، وهذا التوزيع يَصف العديد من الظواهر التي تحدث في الصناعة، والبحوث العلميّة، ومنها على سبيل المثال لا الحصر:

- القياسات الفيزيائية في مجالات عديدة كقياسات الأجزاء الصناعية حيث يكون لتوزيعها الاحتمالي (وفي كثير من الأحيان) تقريب جيد مع التوزيع الطّبيعي.
- الأخطاء في القياسات المعمليّة يمكن تقريبها بشكل جيد من التوزيع الطّبيعي.
- كذلك وجِد أنَّ الكثير من الظواهر الطّبيعيّة تتوزّع احتماليّاً وفقاً لهذا التوزيع أيضاً.

ولهذا السبب أطلق على هذا التوزيع اسم التوزيع الطّبيعي، ويُعرف هذا التوزيع باسم "التوزيع الغاوصي" Gaussian Distribution أيضاً، وذلك نسبةً إلى الرياضياتي الألماني غاوص Johann Carl Friedrich Gauss (1777-1855).

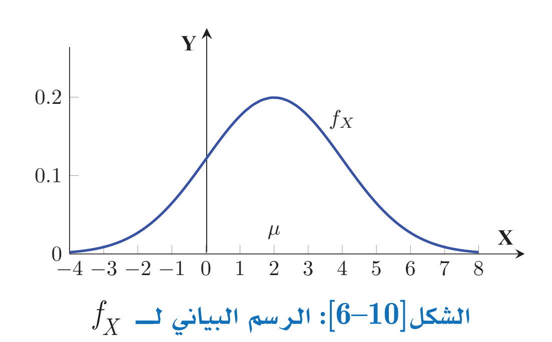
٦-٤-٢- تعريف (المتغيرِّ العشوائيّ الطّبيعي)

Normal Random Variable

يُقال عن متغيرٍّ عشوائيّ X فوق فضاء حوادث ابتدائيّة Ω إنَّه خاضع للتوزيع الطّبيعي بمَعْلَمتين $\mu \in \mathbb{R}$ و $\sigma > 0$ إذا كانت الدالّة التي تميِّزه (ويُرمز لها بـ f_X) مُعطاة من خلال العلاقة الآتية:

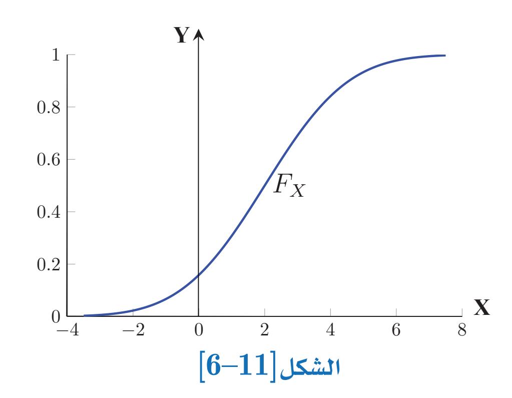
$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \mathbf{exp}\left[-\frac{(x-\mu)^2}{2\,\sigma^2} \right] \quad ; \; x \in \mathbb{R} \qquad [6-13]$$

إنَّ الدالّة f_X تُدعى دالّة الكثافة الاحتماليّة لـ X، ومن أجل $\mu = 2$ و $\sigma = 2$ يكون لرسمها البيانيّ الشكل الآتي (لاحظ أنَّ ذروة الدالّة f_X توافق قيمة μ دائماً):



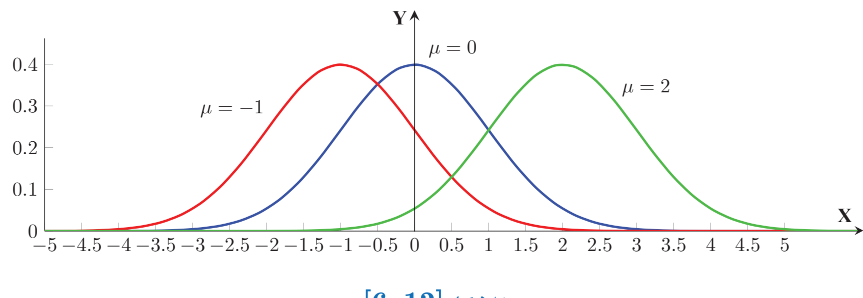
الشكل[6–10]: الرسم البياني لـ f_X

وأمّا الرسم البياني لدالّة التوزيع الطّبيعي ذو المَعْلَمتين $\mu = 2$ و $\sigma = 2$ فله الشكل الآتي:



الشكل[6–11]

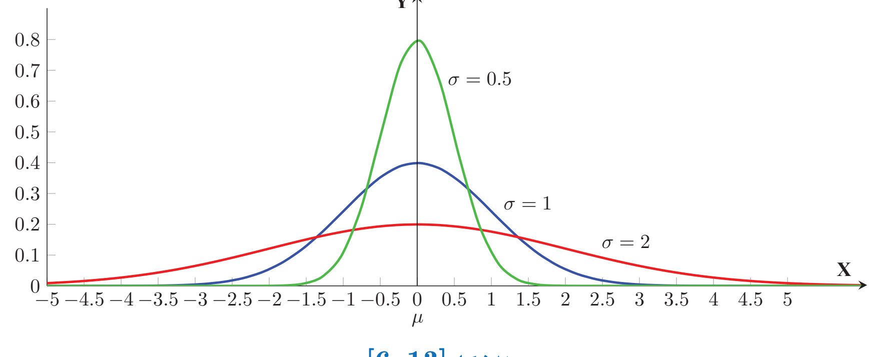
إنَّ الشكل الآتي يقدِّم عرضاً بيانياً لدالّة الكثافة الاحتماليّة لـ X من أجل $\sigma = 1$ وقيم مختلفة لـ μ.



الشكل[6–12]

ويُلاحظ هنا أنَّ تغيّر قيمة μ يؤدّي إلى انسحاب منحني دالّة الكثافته الاحتماليّة f_X يُمنةً أو يُسرى وذلك بحسب القيمة التي تأخذها هذه المَعْلَمة.

الشكل الآتي يقدِّم عرضاً بيانية لدالّة الكثافته الاحتماليّة لـ X من أجل $\mu = 0$ وقيم مختلفة لـ σ.



الشكل[6–13]

نلاحظ هنا أنَّ تغيّر قيمة σ يؤدّي إلى تدبب أو انبساط منحني دالّة الكثافته الاحتماليّة f_X وذلك بحسب القيمة التي تأخذها هذه المَعْلَمة.

٦-٤-٣- ملاحظات

١- إنَّ قيمة المساحة التي تحت منحني دالّة الكثافته الاحتماليّة وفوق المحور **oX** تساوي الواحد تماماً وذلك بغضّ النّظر عن القيم التي تأخذها μ و σ.

٢- يُرمز بـ $\mathbf{N}(\mu , \sigma)$ للتوزيع الطّبيعي ذو المَعْلَمتين μ و σ.

٣- إنَّ قيمة التوقع الرياضياتي والتباين لمتغيِّر عشوائيّ طبيعي X بمَعْلَمتين μ و σ هما على الترتيب $\mathbf{E}(X) = \mu$ و $\mathbf{var}\, X = \sigma^2$، ومن ثمَّ تكون قيمة المَعْلَمة σ هي قيمة الانحراف المعيَّاري لهذا المتغيِّر العشوائيّ.

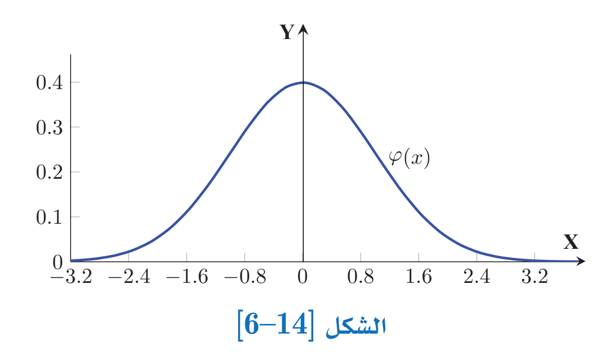
٦-٤-٤- تعريف (المتغيِّر العشوائيّ الطّبيعي المعيّاري)

Standard Normal Random Variable

ليكن X متغيِّراً عشوائياً طبيعياً بمَعْلَمتين $\mu = 0$ و $\sigma = 1$، فعندئذٍ يُقال عن X إنَّه خاضع للتوزيع **الطّبيعي المعيَـاري**، ويكون لدالَّة كثافته الاحتماليّة العرض الآتي:

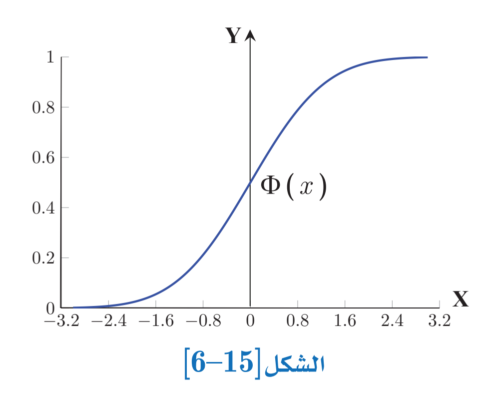
$$f_X(x) = \frac{1}{\sqrt{2\pi}}\, \mathbf{e}^{-\frac{x^2}{2}} \quad ; \quad x \in \mathbb{R} \qquad [6\text{–}14]$$

وفي هذه الحالة الخاصّة يُرمز لدالَّة كثافته الاحتماليّة بـ $\varphi(x)$ بدلاً من f_X، وتكون ذروة منحني دالّة الكثافة الاحتماليّة $\varphi(x)$ واقعةً على المحور **oY** وعلى ارتفاعٍ يساوي 0.339 تقريباً كما يوضِّحه الشكل الآتي:



الشكل [6–14]

وأمّا دالّة التوزيع الطّبيعي المعيّاري فيُرمز لها عادة بـ $\Phi(x)$ بدلاً من F_X، ويكون لرسمها البياني الشكل الآتي:



الشكل[6–15]

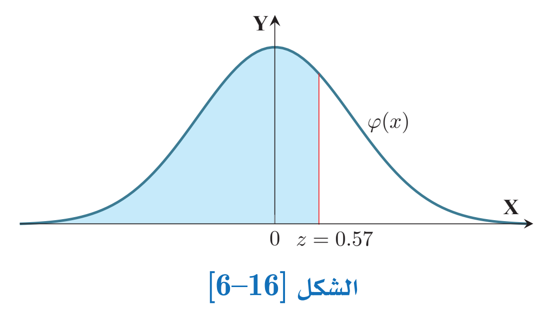
نشير هنا إلى أنَّه يوجد في آخر هذا الكتاب جدول لقيم هذه الدالّة، إحداهما من أجل القيم الموجبة لـ x حيث يبدأ من الصفر وبتزايد يساوي 0.01 ، والأخرى للقيم السالبة لـ x تبدأ من الصفر وبتناقص يساوي 0.01، وهذه القيم تُمكّننا من حساب احتمالات متعلّقة بمتغيّر عشوائيّ طبيعي معيَّاري، فعلى سبيل المثال لو أخذنا Z متغيّراً عشوائياً معيّارياً، فعندئذٍ سيكون لدينا:

$$\Phi(0.57) = \mathbf{P}(Z < 0.57) = 0.7157$$

ونجدها من جدول القيم الموجبة لهذا التوزيع كما يوضِّحها العرض الآتي:

z	**0.00**	**0.01**	**...**	**0.07**	**0.08**	**0.09**
0.0	0.5000	0.5040		0.5279	0.5391	0.5359
⋮						
0.5	0.6915	0.6950		0.7157	0.7190	0.7224
0.6	0.7257	0.7291		0.7486	0.7517	0.7549
0.7	0.7580	0.7611		0.7794	0.7823	0.7852

والتي تمثِّلها المساحة المظللة في الشكل الآتي:



الشكل [6–16]

كما يمكن للقيم التي في الجدول أن تساعدنا في تعيين قيمة z لمتغيِّر عشوائيّ طبيعي معيَّاري Z إذا كانت قيمة الاحتمال $\mathbf{P}(Z < z)$ معلومة، فعلى سبيل المثال لو أخذنا Z متغيِّراً عشوائياً معيَّارياً، وكان لدينا $\mathbf{P}(Z < z) = 0.0021$، فعندئذٍ نجد من جدول التوزيع الطّبيعي المعيَّاري (**القسم الخاص بالقيم السالبة لأنَّ القيمة 0.0021 تقع في جدول القيم السالبة**) أنَّ قيمة z تساوي 2.86-، وذلك من خلال جمع قيمتي z العمودية والأفقية المقابلتين للقيمة 0.0021 وهما 2.8+0.06-، والشكل الآتي يوضِّح لنا ذلك:

z	0.00	0.01	⋯	0.06	⋯	0.09
−3.2	0.0007	0.0007		0.0006		0.0005
⋮						
−2.9	0.0019	0.0018		0.0015		0.0014
−2.8	0.0026	0.0025		0.0021		0.0019
2.7	0.0035	0.0034		0.0029		0.0026

ونشير هنا إلى أنَّه توجد جداول عديدة ومتنوعة في طريقة عرضها والدقة المستخدمة فيها، وقد استخدمنا هذا الجدول لبساطته وسهولة التعامل معه.

٦-٤-٥- ملاحظات

١- من أجل جدول التوزيع الطّبيعي المعيَّاري المقدَّم في آخر هذا الكتاب سوف نضع $\mathbf{P}(Z < z) = 0.9999$ إذا كانت قيمة $3.49 < z$، وأمّا إذا كانت $z < -3.49$ فإنّنا سنضع $\mathbf{P}(Z < z) = 0.0001$.

٢- من أجل أي متغيِّر عشوائيّ مستمر X يكون $\mathbf{P}(X = x) = 0$ مهما يكن $x \in \mathbb{R}$، وكذلك لدينا العلاقتين الآتيتين صحيحتين:

$$\mathbf{P}(X < x) = \mathbf{P}(X \le x) \qquad \& \qquad \mathbf{P}(X > x) = \mathbf{P}(X \ge x)$$

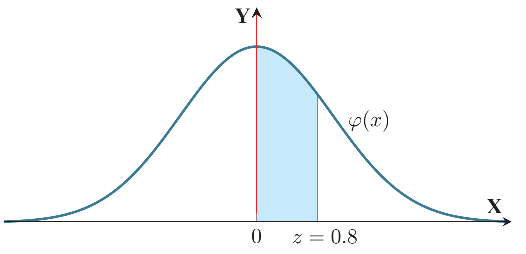
ولذلك فمن أجل متغيِّر عشوائيّ طبيعي معيَّاري Z ستكون لدينا العلاقات الآتية مُحقَّقةً أيضاً:

$$\mathbf{P}(Z = z) = 0 \qquad \& \qquad \mathbf{P}(Z < z) = \mathbf{P}(Z \le z) \qquad \& \qquad \mathbf{P}(Z > z) = \mathbf{P}(Z \ge z)$$

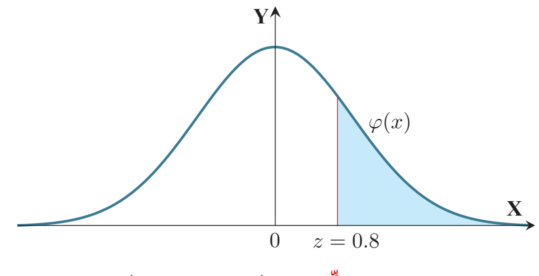
٣- يمكننا تنفيذ نماذج عديدةٍ من الحساب الاحتمالي باستخدام جدول التوزيع الطّبيعي المعيَّاري، ومنها على سبيل المثال الاحتمالات الآتية:

$$\mathbf{P}(Z \ge 0.80) = 1 - \mathbf{P}(Z < 0.80) = 1 - 0.7881 = 0.2119$$

$$\mathbf{P}(0 \le Z < 0.80) = \mathbf{P}(Z < 0.80) - \mathbf{P}(Z < 0) = 0.7881 - 0.5000 = 0.2881$$



المساحة الممثِّلة لـ $\mathbf{P}(0 \le Z < 0.80)$



المساحة الممثِّلة لـ $\mathbf{P}(Z \ge 0.80)$

٦-٤-٥- أمثلة

١- بفرض أنَّ Z متغيِّر عشوائيّ طبيعي معيَّاري، فعندئذٍ لنقم بحساب الاحتمالات الآتية مستخدمين جدول التوزيع الطّبيعي المعياري:

a) $\mathbf{P}(Z < 1.99)$ b) $\mathbf{P}(Z \geq -1.05)$

c) $\mathbf{P}(0.02 < Z < 0.30)$

الأجوبة: لدينا من أجل الفقرة:

a) $\mathbf{P}(Z < 1.99) = 0.9767$

b) $\mathbf{P}(Z \geq -1.05) = 1 - \mathbf{P}(Z < -1.05) = 1 - 0.1469 = 0.8531$

c) $\mathbf{P}(0.02 < Z < 0.30) = \mathbf{P}(Z < 0.30) - \mathbf{P}(Z < 0.02)$

$$= 0.6179 - 0.5080 = 0.1099$$

٢- بفرض أنَّ Z متغيِّر عشوائيّ طبيعيٌّ معيَّاري، فعندئذٍ لنقم بتعيين قيم z التي تحقِّق العلاقات الآتية مستخدمين جدول التوزيع الطّبيعي المعيّاري:

a) $\mathbf{P}(Z < z) = 0.9750$ b) $\mathbf{P}(Z \geq z) = 0.8888$

c) $\mathbf{P}(z < Z < 1) = 0.5557$

الأجوبة: لدينا من أجل:

a) $\mathbf{P}(Z < z) = 0.9750 \;\Rightarrow\; z = 1.96$

b) $\mathbf{P}(Z \geq z) = 0.8888 \;\Rightarrow \left[1 - \mathbf{P}(Z < z)\right] = 0.8888$

$$\Rightarrow \mathbf{P}(Z < z) = 0.1112 \;\Rightarrow\; z = -1.22$$

c) $\mathbf{P}(z < Z < 1) = 0.5570$

$$\Rightarrow \mathbf{P}(Z < 1) - \mathbf{P}(Z < z) = 0.8413 - P(Z < z) = 0.5570$$
$$\Rightarrow \mathbf{P}(Z < z) = 0.8413 - 0.5570 = 0.2843 \Rightarrow z \approx -0.57$$

٦-٤-٦- استعيار المتغيِّرات العشوائيّة Standardization of R.V

إذا كان متغيِّر عشوائيّ X يخضع للتوزيع الطّبيعي غير المعيَّاري، فإنَّه من غير الممكن استخدام جدول التوزيع الطّبيعي المعيّاري لإنجاز حساب الاحتمالات المتعلِّقة بهذا المتغيِّر العشوائيّ بشكل مباشر، ولذلك نقوم بتحويله إلى متغيِّر عشوائيّ طبيعي **معيَّاري** من خلال عمليّة تُدعى استعيَّار المتغيِّر العشوائيّ (**أي تحويل المتغيِّر العشوائيّ إلى متغيِّر عشوائيّ معيَّاري**)، وهذه العمليّة تتمّ على النحو الآتي:

بفرض أنَّ X متغيِّر عشوائيّ له توزيع طبيعي بمَعْلَمتين μ من $\mathbb{R}$ و $\sigma > 0$، وطلب حساب الاحتمال $\mathbf{P}(X < x)$ مع x قيمة مثبّتة من $\mathbb{R}$ فعندئذٍ نضع:

$$Z := \frac{X - \mu}{\sigma} \qquad \& \qquad z := \frac{x - \mu}{\sigma}$$

فيصبح للاحتمال المطلوب العرض الآتي:

$$\mathbf{P}(X < x) = \mathbf{P}\left(\frac{X-\mu}{\sigma} < \frac{x-\mu}{\sigma}\right) = \mathbf{P}(Z < z)$$

حيث تُستَخْرج القيمة $\mathbf{P}(Z < z)$ من جدول التوزيع الطّبيعي المعيَّاري كما فعلنا ذلك سابقاً.

◀ ٦-٤-٧- أمثلة

١- ليكن X متغيِّراً عشوائياً له توزيع طبيعي بمَعْلَمتين $\mu = 15$ و $\sigma = 4$، ولنقم بحساب قيمة الاحتمال $\mathbf{P}(12.5 \le X < 21)$ ، ومن ثمَّ تقديم التفسير الهندسيّ لما يَحدُث نتيجةً لعملية الاستعيَّار.

الحل: من أجل ذلك لنقم أولاً باستعيَّار المتغيِّر العشوائيّ المُعطى X فنجد الآتي:

$$\mathbf{P}(12.5 \le X < 21) = \mathbf{P}\left(\underbrace{\frac{12.5-15}{4}}_{z_1} \le \underbrace{\frac{X-15}{4}}_{Z} < \underbrace{\frac{21-15}{4}}_{z_2}\right)$$

$$= \mathbf{P}\left(-0.625 \le Z < 1.50\right) = \mathbf{P}(Z < 1.50) - \mathbf{P}(Z < -0.625)$$

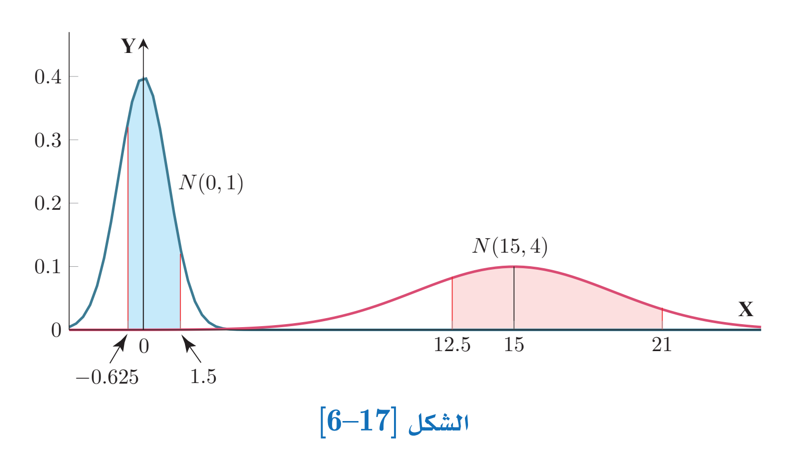
ومن جدول التوزيع الطّبيعي المعيَّاري نجد أنَّ $\mathbf{P}(Z < 1.50) = 0.9332$، وأمَّا لتعيين قيمة الاحتمال $\mathbf{P}(Z < -0.625)$ فإنَّنا نأخذ المتوسط للقيمتين $\mathbf{P}(Z < -0.62)$ و $\mathbf{P}(Z < -0.63)$ لأنَّ القيمة 0.625- غير موجودة في الجدول، ولكنَّها تقع في المنتصف بين القيمتين -0.62 و -0.63، وبالتالي يكون لدينا:

$$\mathbf{P}(-0.625) = \frac{\mathbf{P}(Z < -0.62) + \mathbf{P}(Z < -0.63)}{2} = \frac{0.2676 + 0.2643}{2} = 0.26595$$

ومن ثمَّ يُصبح لدينا:

$$\mathbf{P}(12.5 \le X < 21) = 0.9332 - 0.26595 = 0.66725$$

الآن لو أمعنا النظر في الشكل الآتي لمعرفة ما الذي حدث للمساحة المشغولة من قبل قيمة الاحتمال للمتغيِّر العشوائيّ المُعطى X بعد عملية الاستعيّار، فإنَّنا نلاحظ **تغيُّر شكل المساحة** التي تُمثِّل الاحتمال $\mathbf{P}(12.5 \le X < 21)$ دون تغيُّر قيمتها.



الشكل [6–17]

٢- إذا كانت درجات الاختبار النهائي لطلاب مقرَّر الإحصاء تتبع التوزيع الطّبيعي بمتوسط $\mu = 40$ وانحراف معياري $\sigma = 3$، واختير طالب من الذين تقدموا للاختبار بشكل عشوائيّ، فما هو احتمال:

أ- أن تكون درجته أصغر من 38 درجة؟

ب- أن تكون درجته أكبر من 45 درجة؟

ج- أن تكون درجته ما بين 38 و45 درجة؟

الحل: من أجل ذلك سنأخذ X متغيراً عشوائياً راصداً لدرجة الطالب، فعندئذٍ سيكون لهذا المتغيِّر العشوائيّ توزيع طبيعي $\mathbf{N}(40,3)$، وبالتالي من أجل الطلب:

أ- يكون الاحتمال المطلوب هو $\mathbf{P}(X < 38)$، حيث لدينا:

$$\mathbf{P}(X < 38) = \mathbf{P}\left(\frac{X-40}{3} < \frac{38-40}{3}\right) = \mathbf{P}(Z < -0.67) = 0.2514$$

علماً أنَّنا وضعنا $Z = \dfrac{X-40}{3}$.

ب- يكون الاحتمال المطلوب هو $\mathbf{P}(X > 45)$، حيث لدينا:

$$\mathbf{P}(X > 45) = 1 - \mathbf{P}(X \le 45) = 1 - \mathbf{P}\left(\frac{X-40}{3} \le \frac{45-40}{3}\right)$$
$$= 1 - \mathbf{P}(Z \le 1.67) = 1 - \mathbf{P}(Z < 1.67) = 1 - 0.9525 = 0.0475$$

ج- يكون الاحتمال المطلوب هو $\mathbf{P}(38 \le X \le 45)$، حيث لدينا:

$$\mathbf{P}(38 \le X \le 45) = \mathbf{P}(X < 45) - \mathbf{P}(X \le 38)$$
$$= \mathbf{P}\left(\frac{X-40}{3} < \frac{45-40}{3}\right) - \mathbf{P}\left(\frac{X-40}{3} \le \frac{38-40}{3}\right)$$
$$= \mathbf{P}(Z < 1.67) - \mathbf{P}(Z \le -0.67) \quad = 0.9525 - 0.2514 = 0.7011$$

٣- إذا كانت درجات الحرارة (مقدَّرةً بالدرجات المئوية) لعدّة أصناف من المياه المعدنية لها توزيع طبيعي بمتوسط $\mu = 0$ وانحراف معيّاري $\sigma = 2$ درجة مئويّة، واختير أحد هذه الأصناف عشوائياً، فما هو احتمال:

أ- أن تكون درجة تجمده أعلى من الصفر؟

ب- أن تكون درجة تجمده أدنى من 2- درجة مئوية؟

ج- أن تكون درجة تجمده ما بين 1 و1- درجة مئوية؟

الحل: من أجل ذلك سنأخذ X متغيراً عشوائياً راصداً لدرجة تجمد الماء الذي تمَّ اختياره، فعندئذٍ سيكون لهذا المتغيِّر العشوائيّ توزيع طبيعي $\mathbf{N}(0\,,2)$، وبالتالي من أجل الطلب:

أ- يكون الاحتمال المطلوب هو $\mathbf{P}(X > 0)$، حيث لدينا:

$$\mathbf{P}(X > 0) = 1 - \mathbf{P}\left(\frac{X-0}{2} < \frac{0-0}{2}\right) = 1 - \mathbf{P}(Z < 0) = 1 - 0.5 = 0.5$$

علماً أنَّنا قد وضعنا $Z = \frac{X-0}{2}$.

ب- يكون الاحتمال المطلوب هو $\mathbf{P}(X < -2)$، حيث لدينا:

$$\mathbf{P}(X < -2) = \mathbf{P}\left(\frac{X}{2} < \frac{-2}{2}\right) = \mathbf{P}(Z < -1) = 0.1587$$

ج- يكون الاحتمال المطلوب هو $\mathbf{P}(-1 \le X \le 1)$، حيث لدينا:

$$\mathbf{P}(-1 \le X \le 1) = \mathbf{P}(X < 1) - \mathbf{P}(X \le -1) = \mathbf{P}\left(\frac{X}{2} < \frac{1}{2}\right) - \mathbf{P}\left(\frac{X}{2} \le \frac{-1}{2}\right)$$
$$= \mathbf{P}(Z < 0.5) - \mathbf{P}(Z \le -0.5) = 0.6915 - 0.3085 = 0.3830$$

٤- ليكن X متغيِّراً عشوائياً طبيعياً بمتوسط $\mu = -1$ وانحراف معيّاري $\sigma = 0.5$، فإذا كان من أجل قيمة x من $\mathbb{R}$ لدينا $\mathbf{P}(X < x) = 0.9901$ ، فما هي قيمة x ؟

الحل: من أجل ذلك لنقم باستعيّار المتغيِّر العشوائيّ المُعطى ومن ثمَّ استخدام جدول التوزيع الطّبيعي المعيَّاري لتعيين قيمة x المطلوبة، حيث لدينا:

$$\mathbf{P}(X < x) = \mathbf{P}\left(\underbrace{\frac{X-(-1)}{0.5}}_{Z} < \underbrace{\frac{x-(-1)}{0.5}}_{z}\right) = \mathbf{P}(Z < z) = 0.9901$$

ومن جدول التوزيع الطّبيعي المعيَّاري نجد أنَّ قيمة الاحتمال 0.9901 تقع على تقاطع المسارين للقيمتين 2.3 مع 0.03 ، ومن ثمَّ تكون قيمة $z = 2.33$ ، ومنه تُحسب قيمة x المطلوبة على النحو الآتي:

$$z = \frac{x+1}{0.5} = 2.33 \Rightarrow x = 0.5(2.33) - 1 = 1.165 - 1 = 0.165$$



تمارين



١- لنأخذ تجربة إلقاء حجر نرد متوازن لمرَّة واحدة فقط، ولنُرفق كلّ عدد فرديٍّ من الأعداد التي سنحصل عليها بعددٍ يساوي مجموع الأعداد الفرديّة النّاتجة عن هذه التجربة، وأمّا الأعداد الزوجيّة النّاتجة عن هذه التجربة فسنرفق كلّ واحد منها بضعفه. عندئذٍ المطلوب هو:

أ- هل عملية الإرفاق هذه تعطي متغيرّاً عشوائيّاً، وما هو نوعه ؟

ب- مثِّل هذا المتغيرِّ العشوائيّ جدوليّاً.

ج- مثِّل هذا المتغيرِّ العشوائيّ بيانياً.

د- ما هو احتمال حصولنا على القيمة 9 أو 12 ؟

٢- ليكن Ω فضاء حوادث ابتدائيّة لتجربة عشوائيّة، وX متغيرِّ عشوائيّ على Ω مُعطى بالعلاقة الآتية:

$$\mathbf{P}(X = x) = c(x^2 + 2) \qquad ; x = 0, 1, 2, 3, 4$$

والمطلوب:

أ- عيِّن قيمة الثابت c .

ب- اكتب العلاقات التي تُعطي الاحتمالات الكتلية لهذا المتغيرِّ العشوائيّ.

ج- مثِّل هذا المتغيرِّ العشوائيّ جدولياً.

د- مثِّل هذا المتغيرِّ العشوائيّ بيانياً.

٣- ليكن Ω فضاء الحوادث الابتدائيّة لتجربة إلقاء حجري نرد مُتَمايزين ومتوازنين لمرّةٍ واحدةٍ، وليكن X تطبيقاً حقيقياً معرَّفاً على Ω كما يلي:

$$X\big((i\,,j)\big) := \begin{cases} -1 & for \quad i+j > 10 \\ \ \ 1 & for \quad i+j \le 10 \end{cases} \qquad ; \ i, j = 1,2,3,4,5,6$$

والمطلوب ما يلي:

أ- هل التطبيق X هو متغيرِّ عشوائيّ على Ω.

ب- مثِّل هذا المتغيرِّ العشوائيّ جدولياً.

ج- مثِّل هذا المتغيرِّ العشوائيّ بيانياً.

د- عيِّن دالّة توزيع X وارسمها.

ه- استخدم دالّة توزيع X في حساب الاحتمال $P(-3 \le X < 0)$.

٤- لدى عملية تصنيع منتج من نوع مُحدَّد تقوم لجنة تفتيش عن الجودة بسحبٍ عشوائيٍّ لمجموعات كلٍّ منها مكوَّنة من ثلاثة عناصر منتجة وذلك بغية تفتيشها والتحقّق من أنّها تُلبّي المواصفات المطلوبة، فإذا كنّا ننظر إلى حصولنا على مكوَّن معيب أنّه نجاح Success، وأنَّ لجميع العناصر النصيب نفسه في الاختيار، فعندئذٍ إذا تمَّ سحب ثمان مجموعات وكانت نتائجها كما في الجدول الآتي (**علماً أنّ الحرف S و F في الجدول الآتي يعني النجاح والفشل على الترتيب**):

المجموعة ذات الرقم	8	7	6	5	4	3	2	1
نتائج السحب	FFF	FFS	FSF	SFF	FSS	SSF	SFS	SSS

وبفرض X هو متغيِّر عشوائيّ راصداً لعدد مرّات النجاح في السحوبات جميعاً، فعندئذٍ:

أ- مثِّل هذا المتغيِّر العشوائيّ جدولياً.

ب- مثِّل هذا المتغيِّر العشوائيّ بيانياً.

٥- لنفترض أنَّ متغيِّراً عشوائياً X قُدِّم من خلال الجدول الآتي:

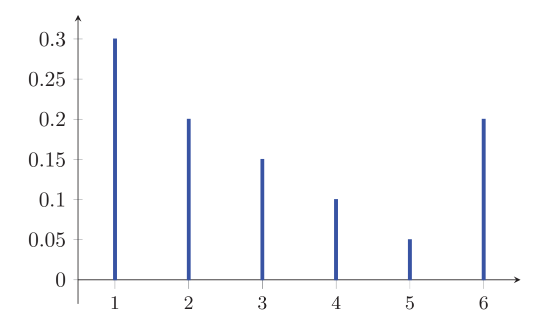
x_i قيم X	0	1	2	3	4	5	Total
$p_i = \mathbf{P}(X = x_i)$	0.15	0.20	؟	0.25	0.07	0.03	1

والمطلوب:

أ- عين القيمة المجهولة في جدول توزيع X .

ج- مثِّل هذا المتغيِّر العشوائيّ بيانياً.

٦- لنفترض أنَّ متغيِّراً عشوائياً X قُدِّم بيانيّاً من خلال الشكل الآتي:



والمطلوب:

أ- عيِّن العلاقات التي تميِّز هذا المتغيِّر العشوائيّ X .

ب- مثِّل هذا المتغيِّر العشوائيّ جدولياً.

٧- يوجد على رف 15 كتاباً منها 4 كتب في الأدب العربي والباقي في موضوعاتٍ أخرى. قمنا بسحب عشوائيّ لـ 3 كتب من تلك التي على الرف، فإذا علمت أنَّ لجميع الكتب النصيب نفسه في السحب (أو الاختيار)، فما هو احتمال:

أ- إذا كان السحب يتمّ على التتالي، فما هو احتمال أن يكون لدينا كتابين في الأدب العربي من تلك التي قمنا بسحبها.

ب- إذا كان السحب يتمّ على دفعةٍ واحدةٍ، فما هو احتمال أن يكون لدينا كتابين على الأقل في الأدب العربي من تلك التي قمنا بسحبها.

٨- قذفت قطعة نقودٍ متوازنةٍ لثلاث مرّاتٍ متتالية، وليكن X متغيِّر عشوائيّ راصد لعدد الشعارات التي سنحصل عليها، والمطلوب:

أ- ما هو احتمال حصولنا على شعارين؟

ب- ما هو احتمال حصولنا على شعارين على الأكثر؟

ج- ما هو احتمال عدم حصولنا على شعار واحد على الأكثر؟

٩- يوجد في مؤسّسةٍ تجاريةٍ 100 عامل منهم 30 أجنبياً. أرادت هذه المؤسسة تشكيل ورشة عمل لتطوير عملها فقامت باختيارٍ عشوائيٍّ لـ 15 عاملاً، عندئذٍ ما هو:

أ- احتمال وجود 10 عمال أجانب في المجموعة المُختارة؟

ب- احتمال وجود 3 عمال أجانب على الأقل في المجموعة المُختارة؟

ج- احتمال عدم وجود أي عامل أجنبي في المجموعة المُختارة؟

١٠- ليكن لدينا X هو متغيّر عشوائيّ مُعطى جدولياً كما يلي:

x_i	1	2	3	4	5	**Total**
$p_i = \mathbf{P}(X = x_i)$	0.15	0.35	0.25	0.13	0.12	**0.15**

والمطلوب حساب التّوقّع الرياضياتي، التّباين والانحراف المعيّاري لهذا المتغيّر العشوائيّ.

١١- بفرض أنَّ باحثاً يقوم بتنفيذ تجربة ما وقد يضطرٌّ إلى تكرارها لمرّات عديدة، فإذا علِمتَ أنَّ هذه التكرارات تتمّ بشكلٍ مستقلٍ كلٍّ منها عن الأخرى، وأنَّ احتمال النجاح في أية تجربة يساوي 0.65، فعندئذٍ:

أ- إذا كان سيكرِّر التجربة حتى الحصول على نجاح لأوَّل مرّة، فما هو احتمال أن يكون قد نفَّذ خمس تجارب؟

ب- إذا قام بتكرار التجربة لسبع مرّات متتالية، فما هو احتمال أن يكون قد نجح في خمس منها على الأكثر؟

١٢- لنقم بإلقاء قطعة نقود متوازنة لتسع مرَّات متتالية، فما هو احتمال الحصول على شعارين على الأقل خلال الرميات التسع؟

١٣- لدينا صندوق يحوي n كرة متماثلة تماماً وسُجِّل على إحداها عبارة رابح، فلو قام شخص ما بعمليات سحب مع الإعادة لكرة من هذا الصندوق وبحيث تُخلط الكرات جيّداً بعد كلّ إعادة، فإذا علمت أنَّ هذا الشخص سيتوقف عن السحب لدى حصوله على الكرة الرابحة، فعندئذٍ:

أ- ما هو احتمال أن يتوقف عن السحب بعد أن يكون عدد الكرات المسحوبة يساوي نصف عدد الكرات التي في الصندوق؟

ب- أعد الطلب السابق إذا كان $n = 50$ ثمَّ من أجل $n = 100$، وماذا تلاحظ.

١٤- في معمل للحاسوب يوجد 25 جهازاً منها ثلاثة أجهزةٍ مُعطَّلة، قام ثلاثة طلاب (وبشكل عشوائيّ) بالجلوس أمام ثلاثة أجهزة كلّ على انفراد ودون معرفتهم بالأجهزة المعطّلة، وذلك من أجل تأدية اختبار لمقرَّر، فما هو احتمال أن يتمكنوا الثلاثة من تأدية اختباراتهم؟

١٥- إذا علمت أنَّ الأخطاء المرتكبة في قياس الطول لقضبانٍ معدنيةٍ من نوعٍ معيّنٍ تتوزَّع طبيعياً معيّارياً، قمنا بسحب عشوائيّ لقضيب من المنتج الكلي، فما هو احتمال أن يكون في طوله زيادة بمقدار 0.75 من وحدة القياس المستخدمة؟

١٦- ليكن X متغيرًّا عشوائياً له توزيع طبيعي بمَعْلَمتين $\mu = 1$ و $\sigma = 0.75$، فإذا كان $-0.25 \le X < 0.75$، فما هما القيمتان اللتان تقع بينهما قيمة المتغيِّر العشوائيّ Z الناتج عن عمليّة الاستعيّار X ؟

١٧- ليكن X متغيِّراً عشوائياً له توزيع طبيعي بمَعْلَمتين $\mu = 1$ و $\sigma = 2$، والمطلوب:

أ- حساب الاحتمال $P(X < 1.53)$.

ب- حساب الاحتمال $P(X > -2.74)$.

ج- حساب الاحتمال $P(-2.74 \le X < 1.53)$.

١٨- ليكن X متغيِّراً عشوائياً له توزيع طبيعي بمَعْلَمتين $\mu = -2$ و $\sigma = 1.5$، والمطلوب:

أ- تعيين قيمة x إذا كان $P(X < x) = 0.9750$.

ب- تعيين قيمة x إذا كان $P(X > x) = 0.8549$.

١٩- إذا كانت درجة غليان ماء الشرب في المنازل (مقدّرة بالدرجات المئوية) تتبع التوزيع الطّبيعي بمتوسط $\mu = 107$ درجة مئوية وبانحراف معياري $\sigma = 3$ درجات مئوية، واُختيِرَت عينّة من خطوط المياه في أحد المنازل، فما هو احتمال:

أ- أن تكون درجة غليان الماء لهذه العينة أدنى من 105 درجات مئوية ؟

ب- أن تكون درجة غليان الماء لهذه العينة أعلى من 110 درجات مئوية؟

ج- أن تكون درجة غليان الماء لهذه العينة ما بين 100 و107 درجات مئوية ؟

المراجع

1. Dimitri P. Bertsekas & John N. Tsitsiklis; Introduction to Probability; Dimitri P. Bertsekas and John N. Tsitsiklis; 2002.

2. Douglas C. Montgomery & George C. Runger; Applied Statistics and Probability for Engineers -Third Edition; John Wiley & Sons, Inc; 2003.

3. Fisz, M.; Wahrschichkeitsrechnung und mathematische Statistik; VEB Deutscher Verlag der Wissenschaften, Berlin; 1980.

4. Gnedenko, B. W.; Lehrbuch der Wahrschinlichkeitsrechnung; Akademie – Verlag. Berlin; 1987.

5. Krause, B. & Metzler, P.; Angewadte Statistik; VEB Deutscher-Verlag der Wessenschaften Berlin; 1988.

6. Kristian, H.; Angewandte Wahrscheinlichkeitsthorie; Friedr, Vieweg and Sohn, verlagsgeselschaft mbH, Braunschweig , Wiesbaden; 2003.

7. Loeve, M.; Probability Theory (I + II); Springer, New York; 1978.

8. Richard J. Larsen & Morris L. Marx; Introduction to Mathematical Statistics and Its Applications - 5/E; Pearson • Cloth, 768 pp; 2012.

9. Ronald E. Walpole & Raymond H. Myers & Sharon L. Myers & Keying Ye; Probability & Statistics for Engineers & Scientists – Ninth Edition; Pearson Education, Inc; 2012.

10. Sirijaev, A.N.; Wahrscheinlichkeit; VEB Deutscher-Verlag der Wessenschaften Berlin; 1988.

11. Trevor Hastie & Robert Tibshirani & Jerome Friedmann; The Elements of Statistical Learning - Second Edition; Springer Science+Business Media, LLC; 2009.

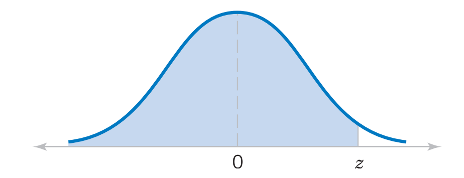
مراجع ومعاجم عربية

١. حميد عويّد العكله و عبيد جفين القحطاني - الاحصاء والاحتمالات (نظرية وتطبيقات) - دار جامعة الملك سعود للنشر - الرياض 2107 (قيد الطباعة).

٢. على مصطفى بن الأشتر- معجم الرياضيات- دار النشر أكاديميا – بيروت 1995.

٣. عدنان، ف.م. - باقر، س.ط. - العايدي، ص.ن. - فران، هـ.ر. و العقيل، ع.س.هـ.- معجم الرياضيات – منشورات مؤسسة الكويت للتقدّم العلمي 1990.

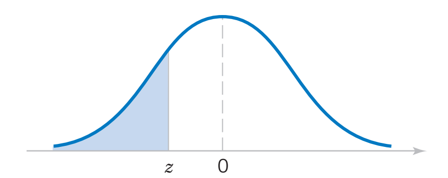
جدول التوزيع الطبيعي المعياري
Standard Normal Distribution Table



من أجل القيم الموجبة لـ Z

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.50 and up	.9999									

جدول التوزيع الطبيعي المعياري
Standard Normal Distribution Table



من أجل القيم السالبة لـ Z

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.50 and lower	.0001									
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

ثبت المصطلحات

A

English	العربية
Absolute Frequency	
Addition Law	قانونّ الجمع (في الاحتمالات)
Additive Rule	قاعدة الجمع (في الحساب)
Almost Certainly Event	حادث شبه أكيد
Almost Impossible Event	حادث شبه مستحيل
Ascending Cumulative Frequency	التكرار المتجمِّع الصاعد
Ascending Cumulative Frequency (Ogive)	مضلَّع التكرار المتجمِّع الصاعد
Associative Property	الخاصة التّجميعيّة

B

English	العربية
Bar Graph	العرض الشرائطي
Bar Notation	طريقة القاعدة (لتعيين مجموعة)
Bays's Formula	صيغة بييز (في الاحتمالات)
Bernoulli Distribution	توزيع برنولي
Bernoulli Experiment	تجربة برنولية
Bernoulli Random Variable	متغيِّر عشوائي برنولي
Bijective	خاصّيّة التقابل (للدوال والتطبيقات)
Bijective Function	دالَّة تقابل
Bijective Map	تطبيق تقابل
Binomial Distribution	التوزيع الحدَّاني
Binomial Random Variable	متغيِّر عشوائي برنولي
Box plot	العرض (أو التمثيل) الصندوقي

C

English	العربية
Cardinality of a Set	قدرة مجموعة
Causal Chain Relation	علاقة السلسلة السببية
Causal Correlation	الارتباط السببي
Celsius	وحدة (الدرجة المئوية) لقياس الحرارة
Central Tendency Measure	مقياس نزعة مركزيَّة
Certain Event	الحادث الأكيد
Class (Group, Category or Interval)	الفئة (في جدول تكراري)
Class Boundaries	الحدود العمليَّة للفئة
Class Limit	الحدود الفعليّة للفئة
Classes Table	جدول تكراري
Classical Definition of Probability	التعريف التقليدي للاحتمال
Coefficient of Determination	مُعامل التحديد

Coefficient of Dispersion	مُعامل التشتت
Coefficient of Skewness	مُعامل الالتواء
Coefficient of Skewness	مُعامل الالتواء
Coefficient of Variation	مُعامل التّغيُّر
Coin	قطعة نقود معدنية
Combinations	التوافيق
Complement of an Event	متمِّم حادث
Comprehensive Survey	المسح الشامل
Commutative Property	الخاصة التّبديليّة (أو التّناظريّة)
Conditional Probability	الاحتمال الشرطي
Constant	ثابت
Constant Function	الدالَّة الثابتة
Continuous	مستمرّ (أو متصل)
Continuous Probability Distribution	توزيع احتمالي مستمرَّ
Continuous Random Variable	متغيّر عشوائي مستمر
Continuous Set	مجموعة مستمرّة (أو متصلة)
Continuous Variable	متغيّر مستمرّ
Correlation	الارتباط
Correlation Coefficients	مُعاملات الارتباط
Countable Set	قابلة للعدّ
Cumulative Distribution Function	دالَّة التوزيع التراكمية
Curve	خط منحن
Curves Fitting	توفيق المنحنيات

D

Degenerate Distribution	التوزيع المضمحل
Degree	درجة
De Morgan's laws	قانونا ديمورغان
Dependent Variable	المتغيِّر التابع
Descending Cumulative Frequency	التكرار المتجمِّع الهابط
Descending Cumulative Frequency Polygon	مضلَّع التكرار المتجمِّع الهابط
Descriptive Measures	مقاييس وصفية
Difference	الفرق (ورمزه \)
Discrete	متقطّع (أو منفصل)
Discrete Probability Distribution	توزيع احتمالي متقطِّع

	Discrete Random Variable	متغيِّر عشوائي متقطِّع
	Discrete Variable	متغيِّر متقطِّع
	Dispersion Measure	مقياس تشتت
	Distributive Property	الخاصة التوزيعية
	Domain	المنطلق (أو مجموعة التّعريف)
	Domain of a Function	المنطلق (أو مجموعة التّعريف) لدالَّة
	Domain of a Map	المنطلق (أو مجموعة التّعريف) لتطبيق
	Dots Representation	التمثيل النقطي
	Draw with Replacement	السحب مع الإعادة (أو الإرجاع)
	Draw without Replacement	السحب بدون إعادة (أو بدون الإرجاع)
E	Element	عنصر
	Elementary Event	حادث ابتدائي
	Empirical Variance	التباين العملي (أو تباين العيِّنة)
	Enumerative Representation	التمثيل السَّردي (لتعيين مجموعة)
	Error	الخطأ
	Existential Quantifier	مُحدِّد الكميّة الوجودي (ورمزه ∃)
	Extreme Value	قيمة متطرفة
F	Fahrenheit	وحدة (فهرنهايت) لقياس الحرارة
	Failure Probability	احتمال الفشـــــــل
	First Quartile (or Lower Quartile Q_1)	الربيعي الأول (أو الربيعي الأدنى)
	Five Number and the Box plot	الأعداد الخمسة والتمثيل الصندوقي للبيانات
	Forecasting	تنبؤ
	Form Measures	مقياس الشكل
	Frequency	تكرار الفئات (عدد العناصر في فئة)
	Frequency Curve	منحني تكراري
	Frequency Distribution Table	جدول توزيع تكراري
	Frequency Histogram	مدرَّج تكراري
	Frequency of Class	تكرار الفئة
	Frequency Polygon	مضلَّع تكراري
	Frequency Table	جدول تكراري
G	Geometric Distribution	التوزيع الهندسي

	English	العربية
	Graphing of Data	عرض البيانات
	Grouped Data	بيانات مُجمَّعة (بيانات مُجدولة)
	Growth Function	دالَّة النمو
H	Head	صورة (يقابلها الكتابة على قطع النقود المعدنية)
	High-Size Sample	عيِّنة كبيرة الحجم
	Hypergeometric Distribution	التوزيع فوق الهندسي (أو الفوهندسي)
I	Image	صورة
	Image of a Function	صورة دالَّة
	Image of a Map	صورة تطبيق
	Impossible Event	الحادث المستحيل
	Independence of Events	استقلال حوادث
	Independent	مستقل
	Independent Events	حوادث مستقلَّة
	Independent Random Variables	متغيِّرات عشوائية مستقلَّة
	Independent Variables	متغيِّرات مستقلَّة
	Initial Element	عنصر البدء
	Injective (or One-to-one)	خاصِّيّة التباين (للدوال والتطبيقات)
	Injective Function (or One-to-one Function)	دالَّة متباينة
	Injective Map (or One-to-one Map)	تطبيق متباين
	Intentional Samples	عيِّنات عمديَّة (أو قصديَّة)
	Interquartile Range (IQR)	المدى الربيعي
	Intersection	التّقاطع (ورمزه $\cap$)
	Interval Scale	مقياس الفترة
	Inverse image of a set	الصورة العكسيّة لمجموعة
K	Kelvin	وحدة (الكلفين) لقياس الحرارة المنخفضة جداً
	Kurtosis	التفلطح (أو التفرطح)
	Kurtosis Measure	مقياس التفلطح
L	Least Square Method	طريقة المربعات الصغرى
	Length of Class	سعة الفئة (في جدول تكراري)
	Leptokurtic Distribution	توزيع مدبب
	Linear	خطيّ

	English	العربية
	Linear Correlation Coefficient	مُعامل الارتباط الخطي
	Linear Function	دالَّة خطيّة
	Lowest Fence	الحاجز الأدنى (أو أدنى حاجز)
M	Map	تطبيق
	Mathematical Expectation	التوقـــــــــع الرياضياتي
	Mathematical Statistics	الإحصاء الرياضياتي
	Mean (or Arithmetic Mean)	المتوسط (أو المتوسط الحسابي)
	Mean of Sample	متوسط العيِّنة
	Median	الوسيط
	Mesokurtic Distribution	توزيع معتدل
	Midpoint	مركز الفئة (في جدول تكراري)
	Mode	المنـوال
	Monetary Units	وحدة نقدية
	Monotone	مطّرد
	Monotony	الاطّراد
	Multimodal Symmetric Distribution	توزيع متناظر متعدّد المنوال
	Multiplication Law	قانون الضرب (في الاحتمالات)
	Multiplication Rule	قاعدة الضرب (في الحساب)
	Mutually Exclusive Events	حوادث متنافية
N	Negative Skewed Distribution	توزيع سالب الالتواء
	Nominal Scale	مقياس اسمي
	Non Linear Correlation	الارتباط غير الخطي
	Non Symmetric Distributions	توزيع تكراري غير متناظر
	Normal (or Gaussian) Distribution	التوزيع الطبيعي (أو التوزيع الغاوصي)
	Number of Classes	عدد الفئات
O	Observation	ملحوظة (أو مشاهدة)
	Observation Values	قيم ملحوظة (أو مشاهدة)
	One Point Distribution	التوزيع وحيد النقطة
	Ordinal Scale	مقياس ترتيبي
	Origin Point	نقطة الأصل (أو مبدأ الإحداثيّات)
	Outcome	نتيجة (نتيجة تجربة)

	Outlier Value	قيمة منعزلة (أو متطرفة)
P	Pairwise Independent	مستقلَّة مثنى مثنى
	Pairwise Mutually Exclusive Events	حوادث متنافية مثنى مثنى
	Parameter	مَعْلَمَة (أو وسيط)
	Partition of a Set	التجزئة لمجموعة
	Percent Frequency	التكرار المئوي
	Percentile (Percentiles)	مئين (المئينات)
	Percentile Frequency Histogram	مدرَّج تكراري مئوي
	Permutation	التباديل
	Person's Coefficient of Correlation	مُعامل بيرسون للارتباط
	Pie Chart	التمثيل بالقطاعات الدائرية (أو بالقرص الدائري)
	Platykurtic Distribution	توزيع منبسط
	Playing Carts	بطاقات لعب
	Polygon	مضلَّع (تمثيل بقطع مستقيمة متتالية)
	Population	مجتمع إحصائي
	Population Parameter	وسيط مجتمع إحصائي
	Position Measure	مقياس موضع
	Positive Skewed Distribution	توزيع موجب الالتواء
	Predicted Value	قيمة مُقدَّرة (أو قيمة متنبأ بها)
	Probability Function	الدالَّة الاحتمالية
	Probability Mass	الكتلة الاحتمالية
	Probability Mass Function	دالَّة الكتلة الاحتمالية
	Probability Theory	نظرية الاحتمالات
	Pull with Return (or with Replacement)	السحب مع الإرجاع (أو الاستبدال)
	Pull without Return (without Replacement)	السحب بدون إرجاع (أو بدون استبدال)
Q	Qualitative Data	بيانات نوعيّة
	Qualitative Variables	متغيِّرات نوعيَّة
	Quantitative Data	بيانات كميّة
	Quantitative Variable	متغيِّر كميّ
	Quartile (Quartiles)	ربيعي (ربيعيات)
R	Random Experiment	تجربة عشوائية

Random Sample	عيّنة عشوائية
Random Variable	متغيِّر عشوائي
Range	المدى (تستخدم من أجل البيانات والدوال والتطبيقات)
Range of a Function	مدى دالَّة
Range of a Map	مدى تطبيق
Rank	رتبة
Ration Scale	مقياس النسبة
Raw Data	بيانات خام
Real Function	دالَّة حقيقيّة
Reciprocal Correlation	الارتباط التبادلي
Reciprocal Relation	علاقة تبادلية
Regression	علم الانحدار
Regression Equation	معادلة انحدار
Regression Line	خط انحدار
Regular Experiment	تجربة نظامية
Relative Frequency	التكرار النسبي
Relative Frequency Histogram	مدرَّج تكراري نسبي
Relative Value	قيمة نسبية
Residual	الباقي
r-Percentile (or rth Percentile)	المئيني الرائي

S

Sample	عيّنة
Sample Function	دالَّة عيّنة
Sampling Distribution	توزيع المعاينة
Sampling Theory	نظرية المعاينة
Scatter Measure	مقياس التبعثر(تسمية أخرى لمقياس التشتت)
Scatter Plot	لوحة الانتشار
Second Quartile	الربيعي الثاني
Set	مجموعة
Set of Outcomes	مجموعة نتائج
Simple Correlation	الارتباط البسيط
Simple Event	حادثٌ بسيط
Simple Function	دالَّة بسيطة
Simple Linear Correlation	الارتباط الخطي البسيط

Simple Random Sample	العيِّنة العشوائية البسيطة
Size of Population	حجم المجتمع
Size of Sample	حجم العيِّنة
Skewed Frequency Distribution	توزيع تكراري ملتوي
Skewed to the Left Distribution	توزيع ملتوي نحو اليسار
Skewed to the Right Distribution	توزيع ملتوي نحو اليمين
Skewness Measure	مقياس التواء
Space of Elementary Events	فضاء الحوادث الابتدائية
Spline	التمهيد الشرائحي
Spread Measures	مقياس الانتشار(تسمية أخرى لمقياس التشتت)
Spurious Correlation	الارتباط الوهمي
Standard Deviation	الانحراف المعياري
Standard Deviation	الانحراف المعيَّاري
Standard Score	درجة معيّارية
Standardization	اسـتعيَّاره (أي جعله معيَّارياً)
Statistic	إحصاء
Statistical Independent	مستقلَّة إحصائياً
Statistical Inference	الاستدلال الإحصائي
Statistics	علم الإحصاء
Steiner Formula	صيغة شتاينر
Stem and Leafs Table	جدول الساق والأوراق
Step Function	دالَّة درجيّة
Stirling Expansion	نشر سترلنغ
Stochastic	علم العشوائيات
Stochastic Independent	مستقلَّة عشوائياً
Straight Regression	مستقيم الانحدار
Stratified Sample	عيِّنة طبقية
Subset	مجموعة جزئية
Success Probability	احتمال النجاح
Surjective	خاصّيّة الغمور (للدوال والتطبيقات)
Surjective Function	دالَّة غامرة
Surjective Map	تطبيق غامر
Symmetric Difference	الفرق التناظري (ورمزه Δ)
Symmetric Frequency Distribution	توزيع تكراري متناظر

	English	العربية
T	Tail	شعار (في تجربة قذف قطعة نقود معدنية)
	Tally	التعداد (تعداد عناصر فئة في جدول تكراري)
	Third Quartile (or Upper Quartile Q_3)	الربيعي الثالث (أو الربيعي الأعلى)
	Total Probability Formula	صيغة الاحتمال التام (أو الكلي)
	Tow Point Distribution	التوزيع ثنائي النقطة
U	Uncountable Set	غير قابلة للعدّ
	Unidirectional Relation	علاقة وحيد الاتجاه
	Unimodal Symmetric Distribution	توزيع متناظر أحادي المنوال
	Union	الاتحاد (أو الاجتماع ورمزه $\cup$)
	Universal Quantifier	مُحدِّد الكميّة الكلي (أو الشّامل، ورمزه $\forall$)
V	Variability Measure	مقياس الاختلاف (تسمية أخرى لمقياس التشتت)
	Variable	متغيِّر
	Variance	التباين
	Variance of Sample	تباين العيِّنة
W	Weighted Mean	المتوسط الموزون
	Whiskers	الشاربان (شاربا التمثيل الصندوقي للبيانات)
Z	Z-Score	الدرجة المعيّارية Z